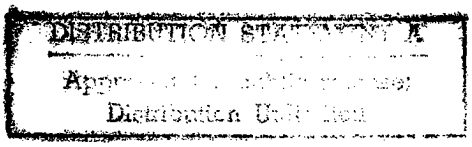
JPRS-CST-92-002
31 JANUARY 1992

JPRS Report

Science & Technology

China

OPTICAL COMPUTING & MEMORY



JPRS-CST-92-002
31 JANUARY 1992

SCIENCE & TECHNOLOGY
CHINA

OPTICAL COMPUTING & MEMORY

CONTENTS

Optical Parallel Multiple Matrix Multiplier

92FE0003A Shanghai ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS] in Chinese Vol 18 No 7, Jul 91 pp 540-544

[Article by Qian Qiuming [6929 4428 2494], Li Qingxiong [2621 1987 3574], Zhao Jianming [6392 1696 2494], and Wang Zhijiang [3769 0037 3068] of Shanghai Institute of Optics and Fine Mechanics (SIOFM) of the Chinese Academy of Sciences, project funded by the NSFC and by the High-Technology Foundation: "Optical Parallel Multiple Matrix Multiplier"; MS received 18 May 89]

[Text] **Abstract**

A parallel matrix multiplier based on only a few imaging systems is presented.

Many computation problems in science and engineering eventually become algebraic operations. Hence, the development of special devices to solve complicated algebraic operations has always been desirable.

This paper presents a parallel matrix multiplier which involves only a few imaging systems. It is based on the corresponding decomposition of optical parallel vector-vector out-product multiplier and matrix multiplier. It can be used to solve real analog multiplication of matrices, analog two-dimensional transforms, digital precision matrix multiplication, and digital precision two-dimensional transforms. It takes full advantage of the parallel nature of optics in all operations. The multiplication of three matrices or a two-dimensional transform takes place within one clock cycle of the system.

I. Structure of the Parallel Matrix Multiplier

1.1 Multiplying Two Matrices

Matrix multiplication can be decomposed into a number of vector inner-products and outer-products (or tensor products).

Figure 1 shows a novel optical parallel matrix multiplier based on vector-vector out-products. The system employs n point light sources on the focal plane of its spherical lens system. In the figure, SL and CL are spherical and cylindrical lenses, respectively. Their focal lengths obey the following equations:

$$f_{SL_1} = f_{CL_1} = f_{CL_2} = f_{CL_4} = f,$$

$$f_{CL_3} = 3f/4$$

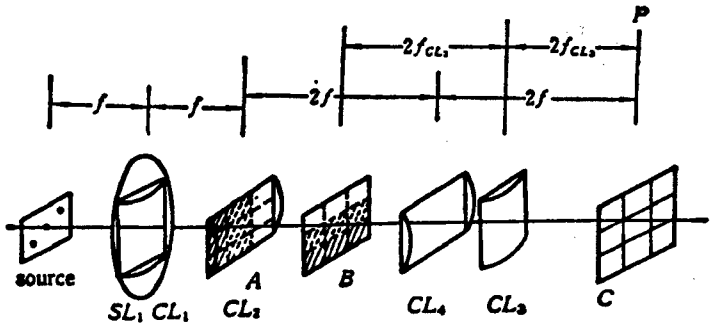


Figure 1. Optical Matrix Multiplier Based on N Vector–Vector Outer Products

In the figure, A and B are spatial light modulators (SLMs) used to input the matrices. They are called input encoding plates. The transmittance of each element is proportional to the value of the corresponding matrix element. Therefore, we have to assume that every matrix element is a positive real number. As for different situations involving numbers that are not positive real numbers, there are a variety of methods available. Even a complex matrix can be coded to become a real matrix before input. In the system shown in Figure 1, addition and multiplication are done in the same manner as with an inner-product matrix multiplier. Figure 1 shows the principle of the optics. Obviously, it is an orthogonal imaging nonspherical optical system.

If the light sources are ideal point sources and all lenses are ideal lenses (without taking aberration into account and cylindrical lenses only form images in one direction), then matrix code plates A and B are placed as shown in Figure 1. If the point source intensities are $I_0\delta(X-X_0)\delta(Y-Y_0)$, $I_1\delta(X-X_1)\delta(Y-Y_1)$, $I_2\delta(X-X_2)\delta(Y-Y_2)$, $\cdots\cdots I_{(n-1)}\delta(X-X_{n-1})\delta(Y-Y_{n-1})$, and the transmittances of encoding plate A and B are (A_{ij}) and (B_{ij}), then it is apparent that there is the following intensity distribution on plane P:

$$C'_{ij} = G\sum_{k=0}^{n-1} A_{ik}B_{kj}I_k$$

If $I_0 = I_1 = \cdots\cdots = I_{n-1} = I$, we have:

$$C_{ij} = C'_{ij}/GI = \sum_{k=0}^{n-1} A_{ik}B_{kj}\cdots\cdots \qquad (1)$$

This is the product of matrix A and matrix B. Of course, transmittance is always less than 1. In reality, it is necessary to multiply by a proportionality factor. As for the rigorous derivation of equation (1), it can be obtained by using a transfer function technique.

1.2 Multiplication of Three Matrices

If we need to multiply three matrices, then the system shown in Figure 2 may be used. the difference is the light source. In Figure 2, n x n point sources (assuming matrix B is n x n) are used. if the three matrices to be multiplied are (A_{ij}), (B_{ij}), and (C_{ij}), then A and C are placed in the manner shown in Figure 2. The light intensity of each light source is proportional to the matrix B_{ij} (i,j = 0, 1, $\cdots\cdots$, n–1), and in their relative positions are also

shown in Figure 2. It is apparent that the light intensity distribution on plane P is:

$$D_{1m} \sim \sum_{i,j=0}^{n-1} A_{1i}B_{ij}C_{jm} \qquad (2)$$

where the symbol "~" represents "proportional to." This is the product of the three matrices. Please note that the order of the three matrices in equation (2) is different from that in Figure 2. This arrangement provides some convenience for the multiplication of more matrices.

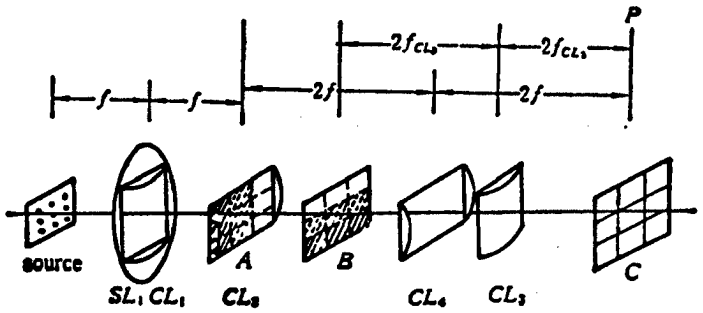


Figure 2. Optical Three-Matrix Multiplier Based on Vector-Vector Outer Product

II. All-Parallel Optical Multiplier Technique and Its Optical Two-Dimensional Transform

Because light intensity is always positive, all matrix elements described in the previous section are positive. In the multiplication of matrices, some elements may be real or complex numbers. This requires us to take certain measures.

This paper presents a technique which involves a combination of polarization encoding and light intensity encoding to realize optical two-dimensional transform.

2.1 Expression and Operation of Intensity/Polarization Hybrid Encoding for Real Numbers

(1) Expressing a real number

To express a real number by intensity and polarization, the magnitude of light intensity can be used to express the absolute value and two different directions of polarization can be used to represent different signs. Based on this concept, an arbitrary real number may be expressed as follows:

$$M = |M|[1 + \text{sign}(M)]/2(+1) + |M|[1 - \text{sign}(M)]/2(-1)$$

where

$$sign(M) = \begin{cases} +1 & M>0 \\ 0 & M=0 \\ -1 & M<0 \end{cases} \qquad (3)$$

Optically, e_x and e_y are two independent orthogonal polarization directions. It is further assumed that e_x corresponds to (+1) and e_y to (-1). Then, the light beam representing the number M has the following two polarization components $\{I_x, I_y\}$:

$$\begin{cases} I_x = |M|[1+sign(M)]/2 \\ I_y = |M|[1-sign(M)]/2 \end{cases} \tag{4}$$

Thus,

$$M = I_x e_x + I_y e_y$$

and

$$I_x I_y = 0 \tag{5}$$

Obviously, this expression is unique. In order to meet the need in practice, the two independent polarization components of the light beam expressing M can be expressed in a more general manner as:

$$\begin{cases} I_x = |M|[1+sign(M)]/2+C \\ I_y = |M|[1-sign(M)]/2+C \end{cases} \tag{6}$$
$$C \geq 0$$

In this case, the expression is not unique.

(2) Arithmetic operation

The multiplication of two real numbers can be expressed by the method shown in equations (4)and (5), i.e.

$$M = I_{Mx} e_x + I_{My} e_y$$
$$N = I_{Nx} e_x + I_{Ny} e_y$$
$$\begin{cases} I_{Mx} = |M|[1+sign(M)]/2 \\ I_{My} = |M|[1-sign(M)]/2 \end{cases}$$
$$\begin{cases} I_{Nx} = |N|[1+sign(N)]/2 \\ I_{Ny} = |N|[1-sign(N)]/2 \end{cases}$$
$$\begin{aligned} MN &= |M||N|[1 + \text{sign}(M)\ \text{sign}(N)]/2e_x \\ &\quad +|M||N|[1 - \text{sign}(M)\ \text{sign}(N)]/2e_y \\ &= |MN|[1 + \text{sign}(MN)]/2e_x + |MN|[1 - \text{sign}(MN)]/2e_y \end{aligned} \tag{7}$$

The multiplication of two real numbers optically expressed by equations (4) and (5) can be realized by shining two polarized light beams on two transmission plates. If the transmittances are η_1, η_2, then

$$I = \eta_1\eta_2[I_{0x}e_x + I_{0y}e_y] \qquad (I_{ox}I_{oy} = 0)$$

In order to express the sign of operation of the two numbers, it is possible to stick a layer of polarizer on the transmission plate. Transmittance and polarization angle together can express a number M with a sign. When $M < 0$, then a layer of polarizer can be used to rotate the direction of polarization of the incident light by 90°. When $M \geq 0$, no polarizer is required and the direction of polarization remains unchanged. In the meantime, $\eta_1 = |M|$. The intensity of the incident beam represents the value of N and its direction of polarization represents the sign of N. Then, the output is (M·N), which in equation (7) is the product of two real numbers.

If addition or subtraction is required, then it is necessary to use the method shown in equation (6). In this case, the expression is no longer unique. For example, the subtraction of two greater than 0 real numbers A and B, i.e., $C = A - B$, can be expressed as:

$$C = I_{Ax}e_x + I_{By}e_y$$

In this case, subtraction becomes adding a positive value to the polarization component which represents the negative value.

Obviously, the number must be transformed after addition or subtraction before multiplication can be carried out again. Nevertheless, it is very difficult to do so directly by optical means. Hence, it is usually done electrically.

2.2 Two-Dimensional Transform of Intensity/Polarization Hybrid Code

Any two-dimensional linear transform in the complex domain can be reduced to four independent real number transforms. Therefore, this section only deals with two-dimensional linear transforms in the real number domain. (This remains true unless specifically expressed otherwise.)

Any two-dimensional transform can be expressed as follows:

$$E = UFV$$

where F is the matrix to be transformed, U and V are the transformation core, and E is the end result. If F_{ij} is expressed by the method described earlier. $|U_{ij}|$ and $|V_{ij}|$ are expressed in terms of encoding plate transmittance and sign (U_{ij}) and sign (V_{ij}) are expressed by the polarizer sticker method. The optical system to realize this transformation is shown in Figure 2. Its light source is:

$$F_{ij} = |F_{ij}|\delta(x-x_i)\delta(y-y_i)\{[1 + \text{sign}(F_{ij})]/2e_x + [1 - \text{sign}(F_{ij})]/2e_y\}$$

The encoding plates are:

$$U_{ij} = |U_{ij}|\delta(x-x_i)\delta(y-y_i)\{[1 + \text{sign}(U_{ij})]/2e_x + [1 - \text{sign}(U_{ij})]/2e_y\}$$
$$V_{ij} = |V_{ij}|\delta(x-x_i)\delta(y-y_i)\{[1 + \text{sign}(V_{ij})]/2e_x + [1 - \text{sign}(V_{ij})]/2e_y\}$$

For a specific type of two-dimensional transform, U_{ij} and V_{ij} remain unchanged. Hence, they can be constructed as fixed encoding plates. Thus, the system does not require an SLM as the input; it only needs two light arrays that can be modulated. If F_{ij} is always greater than 0, it is even simpler. It only needs one light source array, which is extremely convenient.

In Figure 1, all we have to do is to subtract the two polarization components received correspondingly. This only requires equation (6), and the result is apparent. The subtraction of two polarization components can be completed on an electronic circuit. There is no need to do it optically.

In conclusion, the principle of an optical parallel vector-vector outer-product multiplier is presented. It is very simple and can easily be expanded to larger dimensions. By using common optical components, it is possible to obtain the outer-product of $10^4 \times 10^4$ vectors (in analog). The system can be used to realize multiplication involving two and three (positive real number) matrices (in analog). Experimental results to verify this theory will be published separately. (See next issue.)

On the basis of hardware research on an analog matrix multiplier, the algorithm is discussed in this paper. It expands the range of applications to the multiplication of real number matrices. Furthermore, a hybrid spatial intensity/polarization encoding method is used to realize the multiplication of two to three digital (mixed-base) matrices to improve the accuracy of operation. Compared to the two existing all-parallel matrix multipliers (based on inner-product and Fourier transform), this system has a distinct advantage in performing two-dimensional linear transform in the real number domain. It only needs a light array to represent the pattern. The transformation core is a number of fixed encoding plates. The result is obtained in a single step. It does not require an SLM. (This is because matrix B, in the multiplication of matrices A, B, and C, is expressed by a light array in this approach. In other approaches, matrix A is expressed by a light source instead.) Therefore, use of this triple matrix multiplier system should be promoted for two-dimensional transforms.

Experimental Research on Optical Parallel Multiple Matrix Multiplier

92FE0058B Shanghai ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS] in Chinese Vol 18 No 8, Aug 91 pp 606-611

[Article by Qian Qiuming [6929 4428 2494], Li Qingxiong [2621 1987 3574], Zhao Jianming [6392 1696 2494], and Wang Zhijiang [3769 0037 3068] of the CAS Shanghai Institute of Optics and Fine Mechanics (SIOFM), project funded by NSFC and by the Youth Foundation of the Shanghai CAS: "Experimental Research on Optical Parallel Multiple Matrix Multiplier"; MS received 19 Jul 89]

[Text] **Abstract**

The experimental method and results of a parallel optical matrix multiplier are described. The system takes full advantage of optical parallelism and is capable of handling the multiplication of three matrices of a two-dimensional transform within one clock cycle. The sources of error and ways to eliminate them are discussed. Results of multiplications involving two and three matrices (analog operation) are given. The mean error of two-matrix multiplication is 1.2 percent and that of three-matrix multiplication is 1.23 percent.

Currently, matrix operations are done serially in computers. Hence, the inherent parallelism of matrices is totally lost.[1,2] Optical computing is extremely parallel in nature. Therefore, researchers are actively seeking optical means to realize parallel matrix operation. To date, there are many optical matrix multipliers which have been briefly discussed in reference [3]. A parallel matrix multiplier involving only a few imaging elements was introduced in reference [3]. Multiplication of real number matrices is realized based on optical parallel vector-vector outer product and the corresponding decomposition of matrix multiplication. This paper describes the experimental process and results of this parallel matrix multiplier. It makes full use of optical parallelism. The multiplication of three matrices or a two-dimensional transform takes only one clock cycle.

Principle of the All-Parallel Matrix Multiplier

Figure 1 shows the vector-vector outer product based optical parallel matrix multiplier presented in reference [3]. The system employs n x n point sources placed on the focal plane of a spherical lens system for illumination.

In addition, it represents the matrix B. In the figure, SL and CL are spherical and cylindrical lenses, respectively. Their focal lengths follow the equations below:

$$f_{SL_1} = f_{CL_1} = f_{CL_2} = f_{CL_4} = f, f_{CL3} = 3f/4$$

The SLMs used to input the other two matrices are A and C. In our preliminary experiment, they were replaced by input coded plates. (In a two-dimensional transform, these two coded plates represent the core of transformation.) The optical transmittance of each element is proportional to the value of the corresponding matrix element. To this end, we have to assume that all matrix elements are positive real numbers. When they are not positive real numbers, there are a number of schemes available. Even when they are complex numbers, a matrix can be coded to become a positive real number matrix[4] prior to processing. In the system shown in Figure 1, optical addition and multiplication are accomplished in the same manner as calculating the inner product of matrices. Figure 1(a) illustrates the principle of the optical system and Figure 1(b) shows the paths of light rays. Apparently, it is a nonspherical orthogonal imaging optical system.

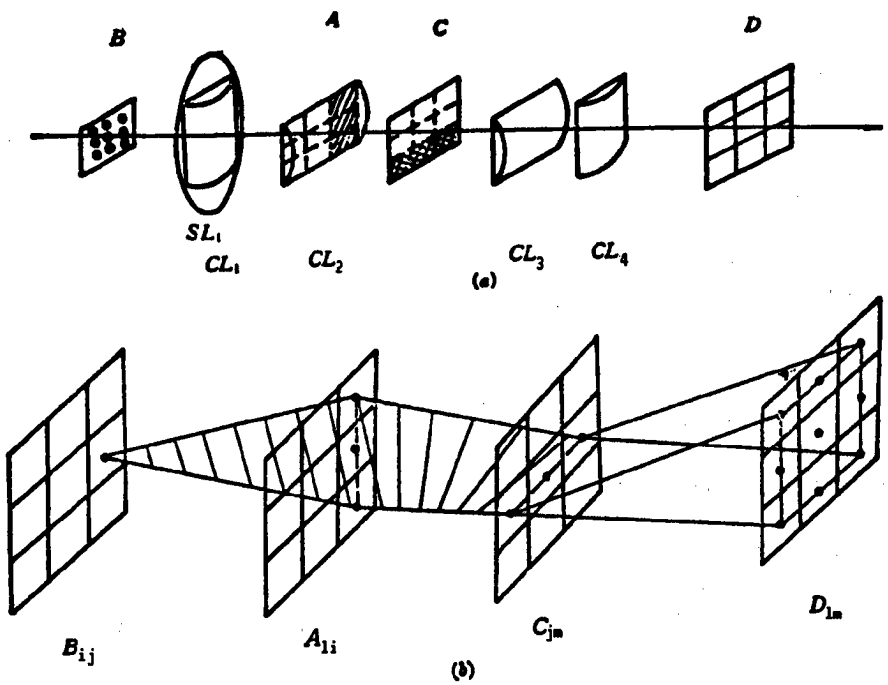


Figure 1. (a) Optical Scheme of the Three-Matrix Multiplication System
(b) Paths of the Light Rays

If the light sources are ideal point sources and all lenses are ideal imaging lenses (without considering any aberrations, and the cylindrical lens forms an image only in one direction), then coded plates A and C are placed according to the way shown in Figure 1(a). If the light intensities of the point sources are B_{1m}, the transmittance values for coded plates A and C are (A_{ij}) and (C_{ij}), then the receiver plane should have the following intensity distribution:

$$D'_{ij} = G\Sigma A_{i1}B_{1m}C_{mj}D_{ij} = D'_{ij}/G = \Sigma A_{i1}B_{1m}C_{mj}\cdots\cdots \quad (1)$$

This is the multiplication of matrices A, B, and C. Of course, transmittance is always less than 1. In reality, it is always necessary to multiply by a proportionality constant.

For the multiplication of two matrices, all we have to do is to change matrix B to a unity matrix, i.e., to change the light source to *n* point sources.

Analysis of the Experimental All-Parallel Optical Matrix Multiplier System

The major sources or error for the matrix multiplier described above include: (1) error due to diffraction of small elements on the coded plates, (2) error due to nonuniform optical angular distribution of different point sources, (3) error from background light, (4) error caused by distortion of coded plates, and (5) error due to deviation of the actual optical system from an ideal system. When the size of the elements of coded plates are large, (2), (3), and (4) have larger impact on the error. It is easy to reduce the last one to meet the accuracy requirement.

1. Relation between size of optical elements and coded plates

Let us assume that the point light sources are s_x and s_y from center to center and the elements on plates A and B are d_{1x}, d_{1y}, d_{2x} and d_{2y} from center to center, respectively, and each element of receiver is a small block h_x x h_y in size. If the system is an ideal optical system, then based on optical wave theory, the resolution light is

$$h_{x1} \approx 1.22\lambda/NA_x$$

$$h_{y1} \approx 1.22\lambda/NA_y$$

Let $h_{x1} = h_{y1} = h_1$

Then $h_x = h_y = h > 2h_1$

It is obvious that

$$S_x = S_y = d_{1x} - d_{2x} = d_{1y} = d_{2y} = h$$

The aperture of the system should be determined by the outer frame of A and B. Hence, there are specific requirements on minimum aperture for every lens.

In order to avoid any overlap of images between neighboring elements on the receiver plane, it is necessary to have an opaque area between neighboring elements. The width of this area is determined by the resolution of each component of the system. Similarly, when each light source is not an ideal point source, the size of the light sources and the gap between them are determined by the resolution of the system and the element size on the encoding plates. Because optical wave theory can only provide a rough approximation with regard to the data (and it often is far off from reality), it will not be discussed in detail. Experimentally, the spacing between two neighboring elements was chosen to be R times the width of the element itself (when approaching the resolution limit of the system, $R > 1$; when far away from the resolution limit, $R \ll 1$).

2. Error due to nonuniform angular light distribution and ways to eliminate it

The error due to nonuniform angular distribution of a light source is obvious. With Gaussian distribution, it can be expressed as:

$$I(\theta) = I_0 e^{-2\theta^2}$$

When the light from a point source passes through the matrix multiplier and reaches A, it can be expressed as follows:

$$I_{AY} \sim e^{-\alpha y^2/f^2}$$
$$y = f \text{ tg } \theta \approx f\theta$$

After penetrating plate A, it is

$$I_{AY} \sim e^{-\alpha y^2/f^2} f_A(y_i)$$

Upon reaching B, it becomes

$$I_{BX} \sim I_{AY} e^{-\alpha x^2/f^2}$$

After penetrating B, it is $I_{BX} \sim I_{AY} e^{-\alpha x^2/f^2} f_B(X_i)$. On the receiver plane, we have

$$I_{xiyi} = Ge^{-\alpha(x^2i+y^2j)/f^2} f_A(y_j) f_B(x_i)$$

where G is dependent upon the diaphragm of the encoding plate.

In order to allow this distribution to meet the accuracy requirement δ, we have

$$f/D = [18a/\ln(\delta+1)]^{1/2}$$

where a is a parameter which defines the Gaussian distribution, δ is the accuracy requirement, and D is the maximum width of the encoding plate, i.e., $D = \max(D_X, D_Y)$. In practice, different light sources may be chosen based on the data. Of course, it is only feasible based on availability.

3. Methods to eliminate error

Spatial distribution multiplication error can be eliminated by a simple method. It involves adding encoding plates with $A_{ij} = C_{ij} = 1$ and $S_{ij} = 1$ onto matrix plates A and C. Use the receiver to measure D'_{ij} and determine $\min(D'_{ij})/D'_{ij} = D_{ij}$. Place an encoding plate with transmittance D_{ij} immediately in front of the receiver. This is a way to eliminate any multiplication error.

Similarly, additive error may be eliminated electronically at the receiver. Using this method, error can be minimized to a very low level (below 1 percent), because most errors in the optical system can be eliminated by these two techniques.

System Adjustment, Experimental Procedure, and Results

Using the system shown in Figure 1, experiments were done to multiply two matrices $A_{10x5}B_{5x10}$ and three matrices U_{10x5}, V_{5x5}, and W_{5x10}.

1. System adjustment

First, align the center of each lens and encoding plates. Adjust each lens so that it is perpendicular to the center line. Then, adjust the two sets of cylindrical lenses to make them orthogonal to each other. Based on the dimension and accuracy requirement of each optical element and encoding plate, the following parameters can be obtained:

$$\text{center deviation} < (d-D)/4;$$

$$\text{vertical deviation} < h/D/2$$

where d is the actual aperture of the optical element, $D = \max(D_x, D_y)$, and h is the minimum width of an encoding plate element.

2. Experimental procedure

In the experiment, f = 10 mm, each encoding plate element is approximately 0.6 x 0.6 mm^2, and the gap is 0.4 mm or 0.3 mm wide. Every cylindrical lens is mounted on a two-dimensionally adjustable rack. Each encoding plate is clamped onto a three-dimensionally adjustable rack. Experimentally, five point sources in B_{ij} were first used for illumination. After using encoding plates $A_{ij} = 1$ and $C_{ij} = 1$, adjustments were made to obtain uniform results on the receiver surface. Then, the real encoding plates were inserted to conduct the tests. The error of the display is 1.0 percent.

3. Experimental results

(1) Multiplication of two matrices:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

A B C

$$ABC = \begin{bmatrix} 3 & 2 & 1 & 1 & 2 & 3 & 2 & 1 & 1 & 2 \\ 2 & 3 & 2 & 1 & 1 & 2 & 3 & 2 & 1 & 1 \\ 1 & 2 & 3 & 2 & 1 & 1 & 2 & 3 & 2 & 1 \\ 1 & 1 & 2 & 3 & 2 & 1 & 1 & 2 & 3 & 2 \\ 2 & 1 & 1 & 2 & 3 & 2 & 1 & 1 & 2 & 3 \\ 3 & 2 & 1 & 1 & 2 & 3 & 2 & 1 & 1 & 2 \\ 2 & 3 & 2 & 1 & 1 & 2 & 3 & 2 & 1 & 1 \\ 1 & 2 & 3 & 2 & 1 & 1 & 2 & 3 & 2 & 1 \\ 1 & 1 & 2 & 3 & 2 & 1 & 1 & 2 & 3 & 2 \\ 2 & 1 & 1 & 2 & 3 & 2 & 1 & 1 & 2 & 3 \end{bmatrix}$$

(2) Multiplication of three matrices:

$$\underset{A}{\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}} \underset{B}{\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}} \underset{C}{\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}}$$

$$ABC = \begin{bmatrix} 3 & 1 & 0 & 1 & 3 & 3 & 1 & 0 & 1 & 3 \\ 4 & 3 & 1 & 1 & 3 & 4 & 3 & 1 & 1 & 3 \\ 3 & 4 & 3 & 1 & 1 & 3 & 4 & 3 & 1 & 1 \\ 1 & 3 & 4 & 3 & 1 & 1 & 3 & 4 & 3 & 1 \\ 1 & 1 & 3 & 4 & 3 & 1 & 1 & 3 & 4 & 3 \\ 3 & 1 & 1 & 3 & 4 & 3 & 1 & 1 & 3 & 4 \\ 4 & 3 & 1 & 1 & 3 & 4 & 3 & 1 & 1 & 3 \\ 3 & 4 & 3 & 1 & 1 & 3 & 4 & 3 & 1 & 1 \\ 1 & 3 & 4 & 3 & 1 & 1 & 3 & 4 & 3 & 1 \\ 1 & 1 & 3 & 4 & 3 & 1 & 1 & 3 & 4 & 3 \end{bmatrix}$$

Figure 2 [photo not reproduced] shows the result of multiplying two matrices. The picture was enlarged (and a correction plate as the one described earlier was inserted in the experiment). For convenience, each element is a plate with either 1.0 or 0.0 transmittance. Table 1 shows the optoelectronic measurement results for the situation shown in Figure 2 [photo not reproduced].

Table 1. Results of Two-Matrix Multiplication (Detected by optoelectrical detector)

3.01	2.02	0.98	1.01	1.98	3.03	2.04	0.99	1.03	2.02
1.97	2.98	1.97	1.02	0.99	1.99	2.99	2.01	1.02	0.98
0.99	2.02	3.03	2.01	1.01	0.98	2.03	3.04	2.03	0.99
1.02	1.01	2.04	3.05	1.97	0.99	1.01	2.03	3.01	1.97
1.98	0.98	1.01	2.03	2.97	1.99	0.99	1.01	1.99	3.03
3.06	2.01	0.99	0.99	2.01	2.97	1.96	0.98	1.01	1.98
2.02	3.01	2.03	1.02	1.01	1.99	2.96	2.02	0.98	0.99
0.99	1.97	2.95	1.97	0.98	0.99	2.03	3.05	2.03	1.01
1.00	0.99	2.01	2.95	2.04	1.01	1.02	2.03	3.03	2.04
1.99	0.97	1.01	2.01	3.02	2.03	1.01	1.00	2.01	3.04

$Error_{max}$=2% $Error_{ave}$=1.2%

References

1. Wang Zhijiang, ZHONGGUO KEXUE YUAN YUANKAN [JOURNAL OF THE CHINESE ACADEMY OF SCIENCES], Vol 2 No 3, 1987, p 198.

2. PROC. IEEE, Vol 72 No 7, 1984, p 1.

3. Qian Qiuming, Li Qingxiong, et al., ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS], Vol 18 No 7, 1991, p 540.

4. H.J. Butterweck, "Principles of Optical Data-Processing," Progress in Optics, edited by E. Wolf, 24, p 213.

Optical Multistability, Multi-Valued Logic Operations Realized Via Composite ZnS Bistable Interference Filters

92FE0058C Shanghai ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS] in Chinese Vol 18 No 8, Aug 91 pp 612-614

[Article by Wang Ruibo [3769 3843 3134], Zhang Lei [1728 7191], Li Chunfei [2621 3196 7378], and Zha Zizhong [2686 1311 1813] of the Department of Applied Physics, Harbin Institute of Technology: "Optical Multistability, Multi-Valued Logic Operations Using Composite ZnS Bistable Interference Filters"; MS received 15 Jun 90]

[Text] **Abstract**

Optical multistability and optical ternary logic gates have been realized using a composite ZnS interference bistable device. The device is an ideal element for studying the feasibility of optical computing because of its favorable characteristics such as ease of fabrication, low threshold power, and room-temperature operation.

Interference-filter-based optical bistable devices are simple to fabricate and easy to operate. They may be used as logic elements in areas such as optical communications, optical computing, and optical information processing. Interference filter-based optical bistable devices have been successfully used in optical full adders,[1] mode recognition,[2] digital conversion,[3] and high-speed switching.[4] This paper reports the use of a composite interference filter bistable device to realize optical multistability and optical symmetric ternary logic gates. Compared to binary logic systems, a multi-valued logic system has higher operating speed and information storage density. Recently, Liu Shutian and M. Watanabe investigated optical multi-valued logic devices such as the Post ternary logic gate and symmetric ternary logic gate using optoelectronic mixers,[5,6] respectively. Their work has considerable significance in improving the capability of optical computing.

I. Optical Multistability

A dielectric interference filter contains alternate coatings of high-refractive index material (such as ZnS or ZnSe) and low-refractive index material (such as Na_3AlF_6 or MgF_2). Each layer is one-quarter wavelength thick.

The film structure is $(HL)^m(nH)(LH)^m$. $(HL)^m$ and $(LH)^m$ are considered as two highly reflective mirrors on both sides and the entire filter is equivalent to a nonlinear F-P standard device. The ZnS interference filter is intrinsically a thermally activated optical bistable device. The bistability is caused by a shift of the center wavelength of the filter to the incident wavelength due to heat. The threshold power to produce bistability is dependent upon the incident angle. The larger the angle of incidence is, the higher the threshold power becomes.

In this experiment, two interference filters were used in series to obtain optical tristability. The experimental setup is shown in Figure 1. The output of the first interference filter is used as the input for the second filter. The filters have a center wavelength of 513.1 nm and 513.8 nm, a FWHM of 2.1 nm and 1.7 nm, and minimum threshold power of 15 mW and 11 mW, respectively. Figure 2 [photo not reproduced] shows a tristability curve obtained by varying the input light power.

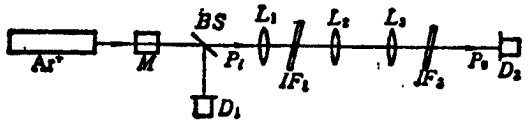


Figure 1. Experimental Setup
Ar^+: argon laser; M: modulator; BS: beam splitter; L_1~L_3: lenses; IF_1, IF_2: interference filters; D_1, D_2: detectors

The tristability produced is in reality the superposition of two bistable effects. As the incident light intensity increases, the output light intensity also continues to rise. When the input light intensity, P_i, is equal to that of the threshold, IF_1, IF_1 is opened. At this time, although IF_2 is still at its lower state, the output light intensity, P_0, still jumps considerably because of its relatively high background transmittance. As the input light intensity increases, when the light intensity passing IF_1 reaches the threshold of IF_2, IF_2 is opened and the output light intensity jumps up once again. This is how tristability is generated. The shape of the curve may be varied by changing the incident angles of IF_1, and IF_2.

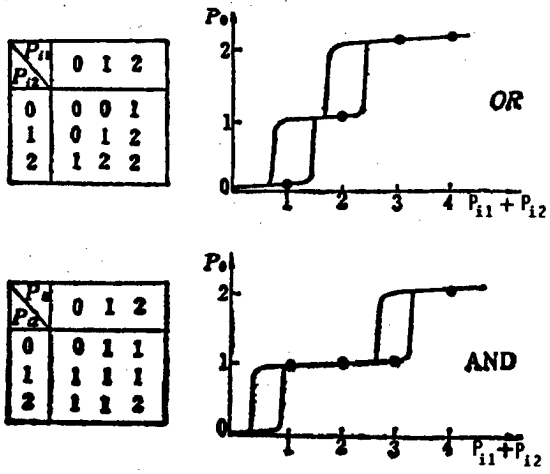


Figure 3. Truth Tables and Input-Output Relation Curves for "OR," "AND" Gates

II. Optical Multi-Valued Logic

The most basic operations in symmetric ternary logic are "AND" and "OR." Figure 3 shows the truth tables and the input-output relation curves for these two logic gates. It indicates that OR and AND gates can be realized by the same optical path. It only requires changing the thresholds of the two bistable devices. Under conditions to meet the logic relation for the AND gate, it is possible to meet the logic relation for the OR gate by raising the threshold of the first bistable device and lowering that of the second device

at the same time. Experimental results of the AND and OR gates obtained using the optical system illustrated in Figure 1 are shown in Figure 4 [photo not reproduced]. Pulses of different amplitudes correspond to different input light intensities. P_0, P_1, and P_2 are the discriminating intensities of 0, 1, and 2, respectively.

The NOT gate [or "inverter"] is also a common logic gate. It may also be realized with two filters, as shown in Figure 5. The light reflected by the second filter is the output. Due to polarization P_b, when the input is 0, 1, and 2, the output is 2, 1, and 0. This meets the logic relation for the NOT gate.

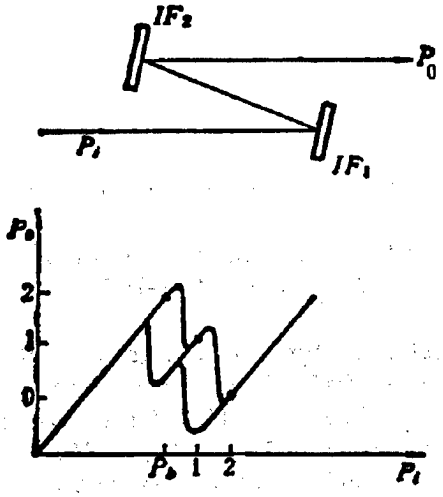


Figure 5. Principle of NOT Operation

III. Discussion

Interference filters are used to study the feasibility of optical computing. Later, they can be replaced by high-speed switches. Because optical systems have a high degree of parallelism, high-processing-speed optical computing systems of practical value can be achieved even with medium-speed logic elements (1 μs~1 ns). The use of medium-speed logic elements is also compatible with available light sources and signal detectors. As progress in interference filter bistable devices advances, ZnSe filters operating in the infrared have been successfully developed. Such devices can be used with laser diodes to substantially reduce the volume of the system.

References

1. B.S. Wherrent, OPT. COMMUN., Vol 56, 1985, p 87.

2. L. Wang, V. Esch, et al., APPL. OPT., Vol 27, 1988, p 1715.

3. L. Wang, et al., SPIE, Vol 752, 1987, p 14.

4. A. Daunois, et al., OPT. COMMUN., Vol 62 No 5, 1987, p 360.

5. Liu Shutian, et al., OPT. LETT., Vol 14, 1989, p 713.

6. M. Watanabe, et al., International Topical Meeting on Optical Computing, 8 April 1990, Japan, paper gC10.

Experimental Study of Optical Parallel Cache Memory Arrays

92FE0089A Shanghai GUANGXUE XUEBAO [ACTA OPTICA SINICA] in Chinese Vol 11 No 9, Sep 91 pp 790-793

[Article by Cao Mingcui [2580 2494 5050], Li Youping [2621 1635 1627], Liu Xiaan [0491 1115 1344], Li Hongpu [2621 3163 6225], and Chen Shaoqian [7115 4801 0107] of the Laser Institute, Huazhong (Central China) University of Science and Technology, Wuhan 430074, project sponsored by the Foundation for Outstanding Young Teachers of the State Education Commission and the National Natural Science Foundation: "Experimental Study of Optical Parallel Cache Memory Arrays"; MS received 16 Feb 90, revised 16 Jul 90]

[Text] **Abstract**

The optical parallel cache memory array and random access array are critical parts of a digital optical processor. A simple optical parallel cache memory array is presented. Furthermore, random access to the optical array has been achieved experimentally.

I. Introduction

Optical mathematical computation is a highly parallel spatial optical logic operation.[1] It allows the simultaneous transfer and processing of a vast amount of optical information through a high-density parallel optical network and two-dimensional optical switching array.[2] From the algorithm and system architecture of digital optical computing, the two-dimensional optical logic array is essential to such a parallel digital optical processor. The optical parallel cache memory array and parallel random access memory array are also vital to such a system.[3] Because digital optical computing is highly parallel, in order to suit high-speed optical processing, the data, command, and program required for processing are first stored in the cache memory array for ready access. In optical information processing, there is also a need to store intermediate results in random access memory to be taken out in the next step.[4]

There are numerous studies on optical bistable logic devices.[5,6] In this work, a random access memory array consisting of two-dimensional optical bistable devices is presented. Details regarding writing, reading, and zero setting are described. Parallel optical random access of information was experimentally achieved.

II. Optical Bistable Memory

In digital optical computing, an optical bistable device can be used as an optical logic switch, as well as optical cache memory array.[7] As an optical logic device, its bistability curve should be adjusted to approach 0, as shown in Figure 1. As an optical memory device, it requires a more stable bistability range in order to implement basic functions such as random writing, storage, reading, and clearing.

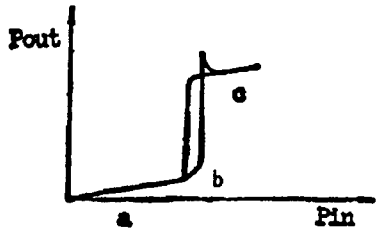


Figure 1. Output/Input Characteristic for Logical Switching

In digital optical computing, the binary code of spatial logic consists of a series of black and white patterns.[1,2] A bright spot has a certain level of power. A dark spot has very low power, approaching 0. For the sake of discussion, let us assume a bright spot is "1" and a dark spot is "0." When a bistable device is used as a memory device, as shown in Figure 2, the power of bias light H_1 is first adjusted to point *b* in the bistable zone. When the input signal is "0," point *b* belongs to the lower state and the output signal is "0." When the input signal is "1," it jumps from point *b* to point *c* which is at a higher state and the output is "1." After the signal is removed, the output still remains at point *b'*. Therefore, the original optical signal "1" is stored in memory. If this information needs to be randomly accessible, an additional optical bistable switch is required. Its characteristic curve is shown in Figure 1. The bias light for this device is set at point *b* before the bistable zone. If one needs to take out the information at a certain point of the memory, the bias light must be applied to the corresponding point on the switch to read it out.

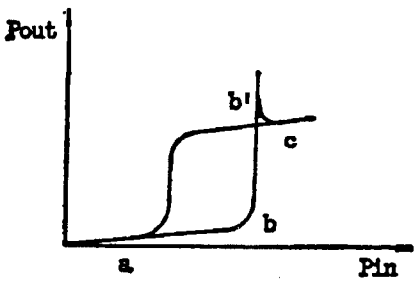


Figure 2. Output/Input Characteristic for the Memory

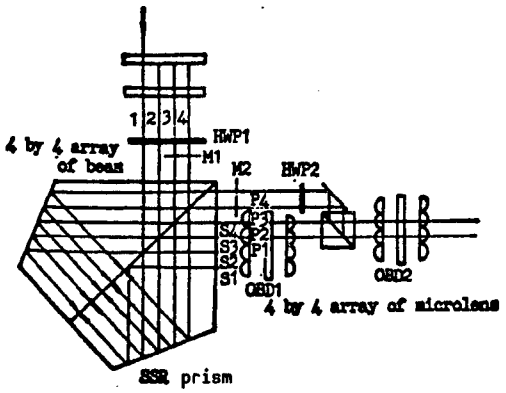


Figure 3. The Experimental Setup of 4 x 2 Array for Parallel Optical Cache Memory Array

If the next piece of information needs to be stored at the same location, then the data stored there must be cleared. This can be done by cutting off the bias light H_1 on the memory and the corresponding bias light H_2 on the switch to set 0.

III. Optical Parallel Cache Memory Array Experiment

Figure 3 shows the experimental setup for the optical system. The optical bistable device is the normal transmission-type optical bistable ZnS device[8,9] developed in our laboratory. Its bistability characteristic curve is shown in Figure 4.

The x and y axes are the input and output power, respectively. In the optical path, the planar thin-film beam splitter splits the beam into a 4 x 4 array. The SSR composite prism and the 4 x 4 microlens array were all developed in our laboratory.

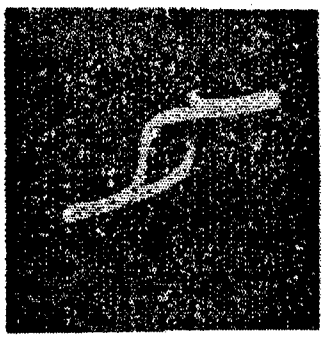


Figure 4. Bistable Characteristic of Normal Transmission ZnS IF Device

The SSR composite lens can divide the 4 x 4 beam array into two arrays of identical intensity. Moreover, the two beam arrays can be displaced relative to each other and make beams from different points in space coincide with one another in a rigorous manner. In this experiment, the first and third column and the second and fourth column were made to overlap by the SSR prism to form a 4 x 2 array, as shown in Figure 3. Using the half wave plate HWP_1, the power of each beam in column P is adjusted to 9 mW. P_1 and P_2 are used as the bias light H_1 for the memory. In the S light array, the S_3S_4 array is the signal light source. Each beam has a power of 3 mW. Arrays S_3S_4 and P_1P_2 are overlapped and passed through the microlens to enter the normal transmission optical bistable device OBD_1 at right incident angle. Array P_3P_4 is used as the bias light H_2 for the second optical bistable device OBD_2. Its optical power may be adjusted by HWP_1.

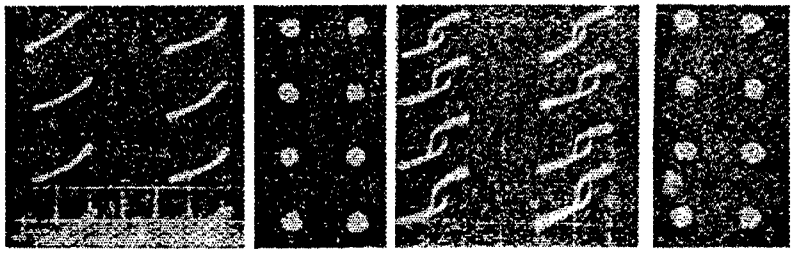


Figure 5. The Process of Writing to, Reading From, and Clearing the Memory in Optical Cache Memory Array

In the experiment, a moving screen M_1 was used to block off light beams in columns 3 and 4. Therefore, P_1P_2 is the only bias light shining on OBD_1. If the bias light is modulated sinusoidally from 0 to 9 mW, the bistable curve of the 4 x 2 array output of OBD_1 is as shown in Figure 5(a). From the photograph, when the bias light power is at 9 mW for every beam, every point of the 4 x 2 array is at its low state. When the input light signal is "0," the OBD_1 output signal is "0." When M_2 is used to block off the light array P_3P_4 and M_1 is removed to let the light array S_3S_4 with optical signal "1" through, every point of the 4 x 2 light input array for OBD_1 simultaneously jumps from "0" to "1," as shown in Figure 5(b). If the incident light (bias and signal) is sinusoidally modulated, i.e., the incident light power is varied from 0 to 12 mW sinusoidally, the bistable curve for every point of the 4 x 2 input array for OBD_1 is as shown in Figure 5(c). The photograph indicates that when the bias is 9 mW and the signal is 3 mW, all points in the 4 x 2 OBD_1 output array are at their high states and the output is "1." Using M_1 to block off S_3S_4, the 4 x 2 output array of OBD_1 is still at the high state, i.e., a bright spot "1" as shown in Figure 5(d). This illustrates that a 4 x 2 input array of "1" is written into and stored in memory. However, there is no light signal output from OBD_2. When the optically stored information is needed, screen M_2 shown in Figure 3 is removed to allow the bias H_2 to

shine on the 4 x 2 array in OBD_2 in the form of array P_3P_4. The stored data can then be read immediately because every point in OBD_2 jumps from a low state to a high state and the output optical signal is "1."

If we want to store the information from the next interval in the same point of memory, the data stored must be cleared first. Before storing new data, the bias light must be removed from the point to be cleared. Afterward, bias light can be restored to begin storing new data. The writing, reading, and clearing of data are shown in Figure 6.

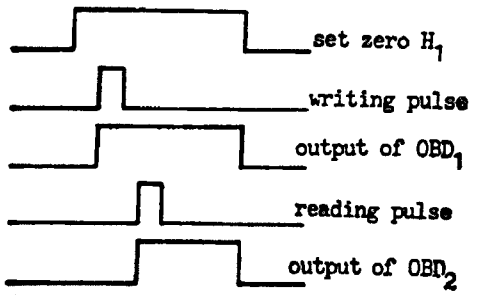


Figure 6. The Process of Writing Into and Reading From the Optical Cache Memory

From this experiment, we know that the system can write, store, read, and clear data. In optical computing, it can serve as a parallel cache memory array, as well as a parallel random access memory.

Conclusions

The experimental optical parallel cache memory array consists of two uniform normal bistable ZnS devices. The optical system includes a high-efficiency planar thin-film beam splitter to provide a uniform beam array. A multipurpose SSR prism ensures the stability of the optical system. This simple optical system can perform basic parallel functions such as write, store, read, and clear in a random fashion. The simultaneous access of an optical signal array is primarily determined by the homogeneity of the beam array and the uniformity of the bistable device. Various parameters of the optical cache memory array depend upon the upward or downward transition time. The reliability of optical storage depends upon the stability of the bistable zone. In digital optical computing, this system can be used as a parallel cache memory array of a parallel random access memory array.

References

1. J.B. Houston, OPT. NEWS, Vol 12 No 4, Apr 1986, p 425.

2. M.J. Murdocca and B. Sugla, APPL. OPT., Vol 28 No 1, Jan 1989, p 182.

3. K.H. Brenner, A. Huang, et al., APPL. OPT., Vol 25 No 18, Sep 1986, p 3054.

4. N. Peyghambarian and H.M. Gibbs, OPT. ENGNG, Vol 24 No 1, Jan 1985, p 68.

5. M.T. Tsao and Liu Xiaan, ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS], Vol 17 No 7, Jul 1990, p 412.

6. A.C. Walker, APPL. OPT., Vol 25 No 10, May 1986, p 1578.

7. M.T. Tsao and Liu Xiaan, GUANGXUE XUEBAO [ACTA OPTICA SINICA], Vol 9 No 12, Dec 1989, p 1129.

8. Liu Xiaan, M.T. Tsao, et al., JIGUANG YU HONGWEI [LASERS AND INFRARED], No 4, 1989, p 27.

All-Optical Matrix Multiplication Using Photorefractive $LiNbO_3$:Fe Crystal

92FE0089B Shanghai GUANGXUE XUEBAO [ACTA OPTICA SINICA] in Chinese Vol 11 No 9, Sep 91 pp 794-800

[Article by Wu Yuanqing [2976 0626 1987], Liu Junmin [0491 6511 3046], Xu Jingjun [6079 0079 0626], Liu Simin [0491 1835 2404], and Zhang Guangyin [1728 0342 1377] of the Physics Department, Nankai University, Tianjin 300071, a project sponsored by the National Natural Science High-Technology Foundation: "All-Optical Matrix Multiplication Using Photorefractive $LiNbO_3$:Fe Crystal"; MS received 6 Aug 90, revised 2 Jan 91]

[Text] **Abstract**

Matrix-vector and matrix-matrix multiplication have been achieved using four-wave mixing in $LiNbO_3$:Fe crystal. Light intensities representing matrix elements are below 100 $\mu W/mm^2$ and the response time is of the order of seconds. The photorefractive effects affecting the accuracy of the operation are also discussed.

I. Introduction

Because all-optical computation based on photorefractive nonlinear optics is parallel and can handle data over a wide frequency band, it is a very attractive subject. Yeh and Chiou[1] first accomplished pixel-pixel multiplication using four-wave frequency mixing in $BaTiO_3$ in 1987. Then, these pixel products were summed with a cylindrical lens to complete the matrix-vector multiplication experiment. This kind of vector-matrix multiplication can be extended to matrix-matrix multiplication by using composite light.[2] Matrix-matrix multiplication can also be done by finding the spatial convolution of four-wave frequency mixing. Chiou and Nguyen[3] implemented this matrix-matrix multiplication scheme with $BaTiO_3$. As far as optical computing is concerned, the light intensity response of the photorefractive crystal must be sufficiently low so that a large number of light beams can be produced by a laser to complete the multiplication of higher-order matrices. Therefore, finding a crystal that has a low response to light intensity is critical. It was experimentally found that the optical response of $LiNbO_3$ with 0.1 wt% (weight percent) Fe is below 100 $\mu W/mm^2$. Furthermore, the photorefractive effects, i.e., light-induced scattering[4] and self-defocusing,[5] have substantial impact on accuracy.

II. Matrix-Vector Multiplication

1. Multiplication by Four-Wave Mixing in Photorefractive Crystal

The four-wave mixing process in a photorefractive crystal can be described by a steady-state wave-coupling equation. As shown in Figure 1, four beams of light with frequency ω and identical polarization intersect in the photorefractive crystal. Let the photoelectric field of these four beams be:

$$E_j(r,t) = A_j(r)\exp[i(k_j\cdot r-\omega t)] + c.c., \tag{1}$$

where $j = 1, 2, 3, 4$. Furthermore, the wave vectors satisfy these relationships: $k_1 = -k_2$, $k_3 = -k_4$. These four beams induce four refractivity gratings in the crystal. Two of them, i.e., transmission grating and reflection grating, are the medium of four-wave mixing. The wave-coupling equations resulting are as follows[6]:

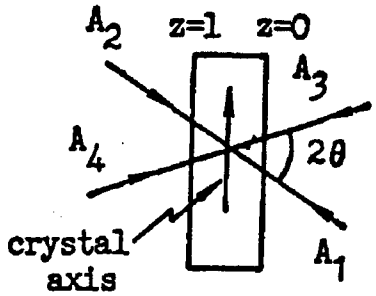


Figure 1. Diagram of the Four-Wave Mixing

(1) The four-wave mixing coupling equations primarily due to the transmission grating are:

$$\left.\begin{aligned}
\frac{dA_1}{dz} &= -\frac{\gamma}{I_0}(A_1A_4^*+A_2^*A_3)A_4-\alpha A_1,\\
\frac{dA_2^*}{dz} &= -\frac{\gamma}{I_0}(A_1A_4^*+A_2^*A_3)A_3^*+\alpha A_2^*,\\
\frac{dA_2}{dz} &= \frac{\gamma}{I_0}(A_1A_4^*+A_2^*A_3)A_2+\alpha A_3^*,\\
\frac{dA_4^*}{dz} &= \frac{\gamma}{I_0}(A_1A_4^*+A_2^*A_3)A_1^*-\alpha A_4^*,\\
\gamma = \frac{i\omega n_1\exp(-i\phi_I)}{2c\cos\theta}&, I_0 = \sum_{i=1}^{4}|A_i|^2,
\end{aligned}\right\} \tag{2}$$

where n_I is the refractivity grating amplitude, ϕ_I is the spatial phase difference between the grating and the interference fringe, c is the speed of light, and θ is the angle between the light velocities.

(2) The four-wave-mixing coupling equations primarily due to the reflection grating are:

$$\left.\begin{aligned}
\frac{dA_1}{dz} &= \frac{\gamma}{I_0}(A_1A_3^*+A_2^*A_4)A_3-\alpha A_1,\\
\frac{dA_2^*}{dz} &= \frac{\gamma}{I_0}(A_1A_3^*+A_2^*A_4)A_4^*+\alpha A_2^*,\\
\frac{dA_3^*}{dz} &= \frac{\gamma}{I_0}(A_1A_3^*+A_2^*A_4)A_1^*+\alpha A_3^*,\\
\frac{dA_4}{dz} &= \frac{\gamma}{I_0}(A_1A_3^*+A_2^*A_4)A_2-\alpha A_4,\\
\gamma &= -\frac{i\omega n_{II}\exp(i\phi_{II})}{2c\cos\theta}.
\end{aligned}\right\} \tag{3}$$

In $LiNbO_3$:Fe crystal, the primary mechanism for electron transport is diffusion. Therefore, $\phi_I = \phi_{II} = (\pi/2)$.[7] Because γ is a real number and when there is almost no absorption ($\alpha = 0$), these two sets of equations can be rigorously solved.[6]

When equation (2), i.e., transmission grating, plays a dominant role, the conjugate reflectivity is

$$R = \left|\frac{A_3(0)}{A_4^*(0)}\right|^2 = \left|\frac{-2cT}{\Delta T + (\Delta^2+4|c|^2)^{1/2}}\right|^2, \tag{4}$$

$$\left.\begin{aligned} &A_3(\ell) = 0,\ I_1+I_4 = const = I_1(0)+I_4(0),\ I_2+I_3 = const = I_2(\ell), \\ &\Delta = (I_2+I_3) - (I_1+I_4) = I_2(\ell) - I_1(0) - I_4(0), \\ &I_0 = I_1 + I_2 + I_3 + I_4 = I_1(0) + I_2(\ell) + I_4(0), \\ &Q = (\Delta^2+4\ |C|^2)^{1/2},\ \mu = (\gamma Q/2I_0),\ T = \tanh\mu\ell, \end{aligned}\right\} \tag{5}$$

$|c|^2$ is determined by the following equation:

$$[|C|^2 - I_1(0)I_2(\ell)]|\Delta T + (\Delta^2+4|C|^2)^{1/2}|^2 + 4|C|^2|T|^2I_4(0)I_2(\ell)$$

$$+ 2|C|^2I_4(0)\ (\Delta^2+4|C|^2)^{1/2}(T+T^*) = 0. \tag{6}$$

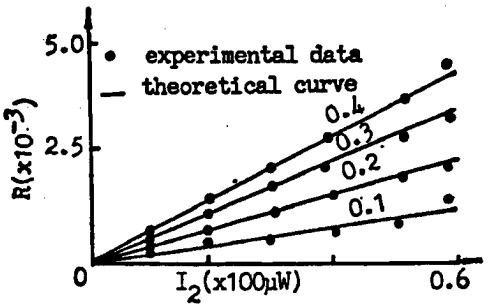


Figure 2. Dependence of the Phase-Conjugate Reflectivity on the Pumping Intensities. The theoretical fitting curves are obtained for $\gamma\ell = -1$, $I_4 = 1.0$ mW. The numerical values above each curve correspond to the intensities of I_1 (x100 μW).

For $I_4 = 1.0$ mW, I_1, $I_2 < 100$ μW, theoretical fitting curves for R versus $I_1(0)$ and $I_2(\ell)$ have been computed and are shown in Figure 2. When $I_4 \gg I_1$, I_2 and I_3, we have

$$\Delta \doteq -I_4(0),\ Q \doteq I_4(0),\ \mu \doteq (\gamma/2), I_0 \doteq I_4(0), \tag{7}$$

Equation (6) can then be simplified as

$$|C|^2 = [(1-T)/(1+T)]^2I_1(0)I_2(\ell). \tag{8}$$

Furthermore, from equation (4), we get

$$R \doteq \left(\frac{1-\exp(-\gamma\ell)}{I_4(0)\exp(\gamma\ell)}\right)^2 I_1(0)\,I_2(\ell) \propto I_1(0)\,I_2(\ell). \tag{9}$$

When equation (3), i.e., reflection grating, is dominant, its conjugate reflectivity is

$$R = \frac{I_2(\ell)}{I_4(0)}\tanh^2[|g_0|\mu(\ell)],$$
$$\tanh^2[|g_0|\mu(\ell)] = \frac{I_1(0)\,I_4(0)\,[1-\exp(-\gamma\ell)]^2}{\{I_1(0)+[I_2(\ell)+I_4(0)]\exp(-2\ell)\}^2+I_1(0)\,I_4(0)\,[1-\exp(-\gamma\ell)]^2}, \quad (10)$$

For $LiNbO_3$:Fe, $|\gamma\ell| \sim 1$. Thus,

$$\tanh^2[|g_0|\mu(\ell)] \doteq \frac{I_1(0)\,[1-\exp(-\gamma\ell)]^2}{I_4(0)\exp(-2\gamma\ell)},$$
$$R = \left[\frac{1-\exp(-\gamma\ell)}{I_4(0)\exp(-\gamma\ell)}\right]^2 I_1(0)\,I_2(\ell) \propto I_1(0)\,I_2(\ell). \quad (11)$$

From equations (9) and (11), we can see that either transmission-grating- or reflection-grating-induced four-wave mixing can accomplish multiplication as long as the probe ratio $q = \{I_4(0)[I_1(0)+I_2(\ell)]\} \gg 1$. When $\gamma\ell < 0$, the transmission grating plays the major role. However, when $\gamma\ell > 0$, the reflection grating dominates. The use of $\gamma\ell$ is determined by the crystal axis and polarization of the coupled light (i.e., the effective photoelectric coefficient γ_{eff}[8,9]).

2. Matrix-Vector Multiplication by Image-Beam/Image-Beam Multiplication

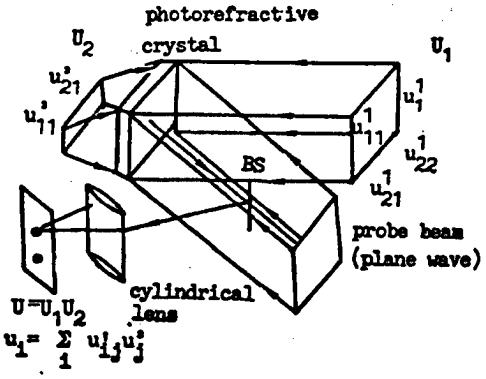


Figure 3. Conceptual Diagram of Matrix-Vector Multiplication

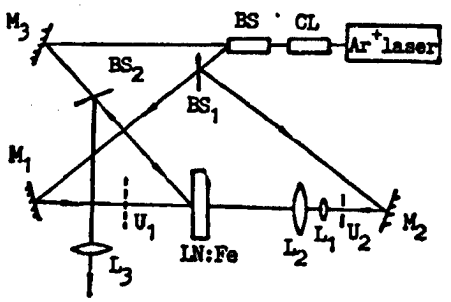


Figure 4. Experimental Setup for Matrix-Vector Multiplication
M_1, M_2, M_3 are mirrors; L_1, L_2, L_3 are cylindrical lenses; BS, BS_1, BS_2 are beam splitters. CL is the collimator.

Figure 3 shows the principle of matrix-vector multiplication. The matrix and vector elements are coded by two aperture masks with light passing in opposite directions. A collimated intense beam of light is used as the probe. Experimentally, a 1.5-mm-thick $LiNbO_3$:Fe crystal is used for the image-beam/image-beam multiplication. These products are "read" by the probe light, which is focused by a cylindrical lens. The summation is done on the focal plane to complete the matrix-vector multiplication process. The experimental setup is shown in Figure 4. Figure 5 shows the result of the product of a 2 x 2 matrix and a 2 x 1 vector. For simplicity, a thick beam was used in the experiment, instead of using cylindrical lenses L_1 and L_2 to expand the beam.

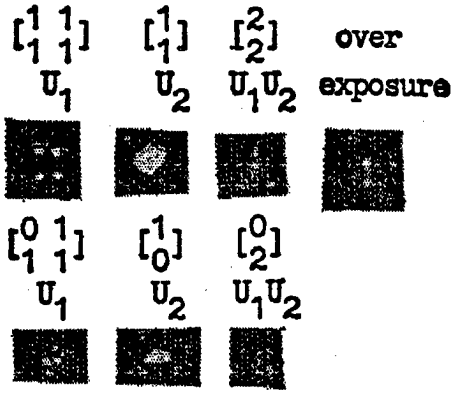


Figure 5. The Operation of Matrix–Vector Multiplication

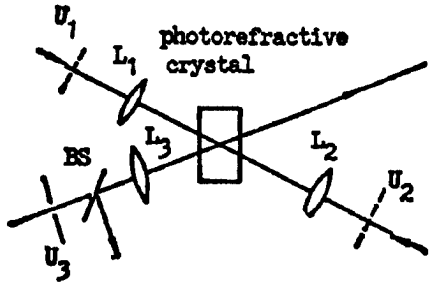


Figure 6. Conceptual Diagram of Matrix Multiplication Using Convolution

III. Matrix-Matrix Multiplication

1. Convolution and Correlation Using Four-Wave Mixing

White and Yariv proved that the correlation and spatial convolution of two patterns could be done by four–wave mixing.[10] The crystal used in that experiment is $Bi_{12}SiO_{20}$ and its principle is shown in Figure 6. After Fourier transformation, the object plane input amplitude $U_1(x,y)$, $U_2(x,y)$, and $U_3(x,y)$ become u_1, u_2, and u_3 ($u_3 \gg u_1$, u_2) on the confocal plane. These three beams are coupled in the crystal to produce four–wave mixing. The conjugate of u_3, u_4, is converted to $U_4(x,y)$ by inverse Fourier transformation using lens L_3. It can be expressed as follows[10]:

$$U_4(x,y) = \psi U_3(x,y) \bigstar [U_1(x,y) \bullet U_2(x,y)], \tag{12}$$

where ψ is a transformation constant, $\bigstar$ and $\bullet$ are correlation and convolution, respectively.

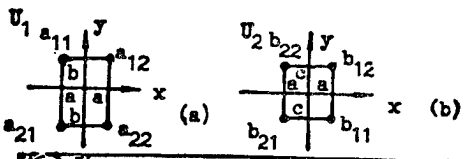
When U_1 and U_2 are the matrix masks and U_3 is an aperture (δ function), we have

$$\begin{aligned} C(x,y) &= U_1(x,y) \bullet U_2(x,y) = \int_{-\infty}^{\infty}\int U_1(\xi,\eta)\, U_2(x-\xi, y-\eta)\, d\xi\, d\eta \\ &= \sum_{ij,i'j'} a_{ij} b_{i'j'} \delta\left(x-\xi^1_{ij}-\xi^2_{i'j'},\ y-\eta^1_{ij}-\eta^2_{i'j'}\right), \end{aligned} \tag{13}$$

$$\begin{aligned} U_4(x,y) &= \psi U_3(x,y) \bigstar [U_1(x,y) \bullet U_2(x,y)] \\ &= \psi \int_{-\infty}^{\infty}\int U_3(\xi,\eta)\, C(x+\xi, y+\eta)\, d\xi\, d\eta \\ &= \psi \int_{-\infty}^{\infty}\int \delta(\xi,\eta)\, C(x+\xi, y+\eta)\, d\xi\, d\eta \\ &= \psi C(x,y), \end{aligned} \tag{14}$$

where $(\xi_{ij}^{Y}\ \eta_{ij}^{2})$ is the coordinate of element ij on matrix mask U_r. Equation (14) shows that after passing through lens L_3, the output amplitude of the conjugate wave produced by four-wave mixing on the focal plane is the convolution of U_1 and U_2.

If U_1 and U_2 are the matrices shown in Figure 7(a) and (b), then their convolution is as shown in Figure 7(c). The center column of the convolution corresponds to the elements of U_1U_2.



x / y	-2a	0	2a
-(b+c)	$a_{21}b_{21}$	$a_{21}b_{11}+a_{22}b_{21}$	$a_{22}b_{11}$
c-b	$a_{21}b_{22}$	$a_{22}b_{22}+a_{21}b_{12}$	$a_{22}b_{12}$
b-c	$a_{11}b_{21}$	$a_{11}b_{11}+a_{12}b_{21}$	$a_{12}b_{11}$
b+c	$a_{11}b_{22}$	$a_{11}b_{12}+a_{12}b_{22}$	$a_{12}b_{12}$

(c)

Figure 7. The Convolution of U_1 and U_2

2. Matrix-Matrix Multiplication by Spatial Convolution

Figure 8 shows the experimental setup. The horizontal distance of holes on U_1 and U_2 is 2a = 1.5 mm, and the vertical distances are 2b = 3 mm and 2c = 1.5 mm. Figure 9(a)~(c) shows the experimental results of matrix multiplication. [The cross terms of the convolution (left and right column in Figure 7(c)) are not shown because they have no contribution to the matrix multiplication.]

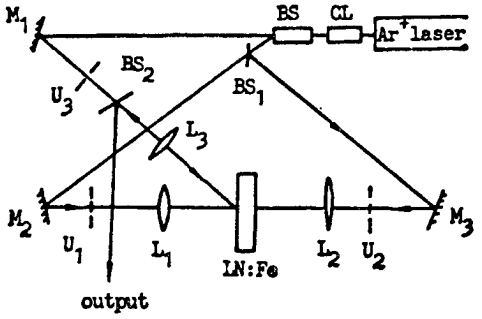


Figure 8. Experimental Setup of Matrix-Matrix Multiplication
M_1, M_2, M_3 are mirrors; L_1, L_2, L_3 are spherical lenses with equal focuses; U_1, U_2, U_3 are matrix masks

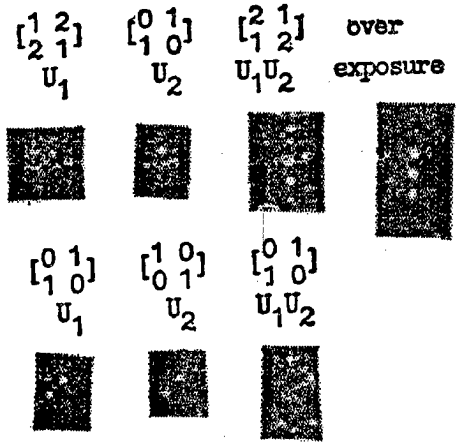


Figure 9. The Operation of Matrix-Matrix Multiplication

IV. Experimental Results and Discussion

Major factors affecting accuracy are self-defocusing and light-induced scattering. Self-defocusing diffuses the output light, as shown in Figure 5. For long exposure, the light spots become undistinguishable. Self-defocusing is caused by an index change induced by a Gaussian beam, as shown in Figure 10(b). If a light beam is uniform and weak, then the index of refraction change is flat and small. Consequently, the self-defocusing effect is negligible. Light-induced scattering causes the conjugate wave to lose some energy.

As shown in Figure 10(a), multiplication is accurate in the stable region marked in the figure. If scattering is strong enough, this stable region will disappear. Light-induced scattering is dependent upon the diameter and intensity of the beam. Zhang Guangyin[11] pointed out that there is a threshold diameter for this type of scattering. Therefore, a low-intensity and small-diameter beam can significantly reduce the effect of light-induced scattering and self-defocusing. In this experiment, the image beam intensity is less than 100 $\mu W/mm^2$ and the diameter is less than 0.1 mm. A stable area of more than two seconds was obtained experimentally. In addition, in matrix-matrix multiplication by convolution, long-time exposure will lead to noise diffraction spots, as shown in Figure 9. These noise diffraction spots are higher-order diffraction points of the transmission or reflection grating. Hence, whether for matrix-vector or matrix-matrix multiplication, there is a definite time within which the result is accurate.

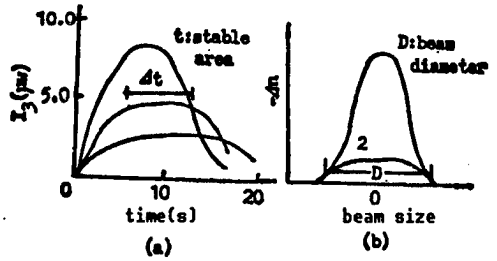


Figure 10. Influences of Self-Defocusing and Light-Induced Scattering
(a) The energy loss caused by scattering.
(b) The refraction index change caused by laser beam. 1. Index change caused by Gaussian beam. 2. Index change caused by weak uniform beam.

V. Conclusions

All-optical computation using photorefractive nonlinear optics has a wide range of future applications. In order to make it practical, the light intensity response of photorefractive materials must be significantly reduced. Experimentally, it was determined that $LiNbO_3$ crystal heavily doped with Fe (0.1 wt%) has a great deal of potential in this field.

The authors wish to thank Associate Professor Zhang Chunping [1728 2504 1627] and Instructors Li Laobing [2621 5071 0365] and Cui Yufen [1805 3768 5358] for their assistance in the experimental work.

References

1. P. Yeh and A.E. Chiou, OPT. LETT., Vol 12 No 2, Feb 1987, p 138.

2. A.E. Chiou, M. Khoshnevisan, et al., PROC. SPIE, Vol 881, 1988, p 250.

3. P. Yeh, et al., OPT. ENGNG., Vol 28 No 4, Apr 1989, p 328.

4. E.M. Avakyan, K.G. Belabaev, et al., SOV. PHYS. SOLID STATE, Vol 25 No 11, Nov 1983, p 1887.

5. Liu Simin, Zhang Guangyin, et al., WULI XUEBAO [ACTA PHYSICA SINICA], Vol 37 No 2, Feb 1988, p 268.

6. Mark Cronin-Golomb, et al., IEEE J. QUANT. ELECTRON., Vol QE-20 No 1, Jan 1984, p 12.

7. D.L. Staebler and J.J. Amodei, J. APPL. PHYS., Vol 43 No 3, Mar 1972, p 1042.

8. Kenneth R. MacDonald, et al., J.O.S.A., Vol 73 No 5, May 1983, p 548.

9. Jack Feinberg, et al., APPL. PHYS., Vol 51 No 3, Mar 1980, p 1297.

10. J.O. White and Amnon Yariv, APPL. PHYS. LETT., Vol 37 No 1, Jul 1980, p 5.

11. Gangyin Chang, et al., APPL. OPT., Vol 25 No 17, Sep 1986, p 2955.

Inter-Pattern-Association Memory With Fresnel Hologram

92FE0089C Shanghai GUANGXUE XUEBAO [ACTA OPTICA SINICA] in Chinese Vol 11 No 9, Sep 91 pp 801-804

[Article by Liu Mingzhe [0491 2494 0772], Wang Xuming [3769 6079 2494], and Mu Guoguang [3018 0948 0342] of the Institute of Modern Optics, Nankai University, Tianjin 300071, and Zhan Yuanling [2069 0337 7881] of the Department of Physics, Nankai University, a project funded by the National Natural Science Foundation: "Inter-Pattern-Association Memory With Fresnel Hologram"; MS received 8 Oct 90, revised 7 Dec 90]

[Text] **Abstract**

An IPA (Inter-Pattern Association) memory based on Fresnel hologram is presented. It has higher recognition capability for nonindependent storage modes. Using Fresnel holograms, an optical IPA memory with large number of neural elements can be implemented.

I. Introduction

Neural networks have been an active area of research in recent years. Associative memory is the most critical component in a neural network. In most optical neural networks, Hopfield's model is used to form associative memory. However, Hopfield's model does not consider the relation between various modes in storage. It only stresses the inter-pattern association of elements. Hopfield's model is valid when storage modes are independent and the ON and OFF probability for each element is equal. Such a system is a poor model for dependent modes. When it is necessary to distinguish similar modes, the difference between them becomes very important. Therefore, when we develop an IPA memory, the similarities and dissimilarities of storage modes must be considered to make them function differently in the association process.

In 1990, Taiwei Lu[1] proposed using an IPA model to fabricate an associative memory. An optical neural network, consisting of a monitor, lens array, spatial light modulator, and computer, has been constructed. The IPA associative memory matrix and the effect of the lens array on each pixel are shown on the monitor. The results indicate that the address-finding capability of the IPA neural network is substantially higher than that of the Hopfield's model.

Especially when the Hamming distance of the storage modes is small, it is capable of finding the address correctly. Thus, the IPA neural network has more stable capability than that of the Hopfield's model. However, because of the resolution of the monitor, the number of pixels stored in the system cannot be very large.

A hologram associative memory is particularly suitable for the storage of two-dimensional pictures with multiple pixels. A theoretical analysis of the characteristics of Fresnel holograms is presented in this paper. An IPA associative memory was fabricated on the basis of this analysis. Furthermore, a large-pixel IPA optical associative memory has to be implemented.

II. Theoretical Analysis

The IPA model decomposes every mode into several parts based on how one mode coincides with other modes in the set of storage modes. Each part has a different function in the associative process. A unique part of the mode has a more important function than other common parts in association.

For example, in storing three modes, when some part of the input address-finding mode is ON, if the corresponding parts of the three modes are all ON, then the input mode may be any one of the three modes. There is no way to tell which one; we can only say there is an input mode. Hence, this part only stimulates the common parts. If the corresponding parts of two modes are ON and that of the third mode is OFF, then the input mode is one of the former two and it is not the latter. However, it is not possible to distinguish between the two. Then, this part stimulates the common part of the former two, but suppresses a unique part of the latter. If only the corresponding part of one mode is ON and those of the other two are OFF, then the input must be the former. This part stimulates all the parts of this mode and suppresses some part of the other two modes. Taiwei Lu proposed a triple gray-scale associative matrix. The elements are -1, 0, and 1. In an optical neural network, it is easy to implement stimulation and suppression. This kind of associative matrix is relatively easy to materialize optically. Nevertheless, the function of each part in the associative process in an IPA memory is proportional to the number of pixels contained in this part. When the Hamming distance of two storage modes is small, there is a substantial overlap between modes. Their unique parts play a limited role in the associative process. It is more difficult to resolve these modes. We believe that we should add a gray scale in the associative matrix to balance different parts. Since the system chooses a threshold for output, the gray-scale requirement is not very stringent, which is easier to realize optically.

The following is a theoretical analysis of IPA memory with Fresnel hologram. The optical path used to record the Fresnel hologram is shown in Figure 1. P_2 is the imaging plane of plane P_1 through lens L_1. P_3 is the imaging plane of plane P_4 through lens L_2. During recording, a transparent plate with a transmissivity of $a(p,q)$ is placed on plane P_2 to come in contact with the holographic mask. It is illuminated by another focused light. Another transparent plate with a transmissivity of $b(x,y)$ is placed on plane P_4. A diffuse scattering screen is placed in front of it.

Let us assume that the diffraction field of b(x,y) on plane P_2 is B(p,q). In addition, the distribution of this field on plane P_3 is noted to be D(x', y') $b^{1/2}$(x', y'). Then, we have

$$B(p,q) = \iint D(x',y')\, b^{1/2}(x',y') \exp[-ik(d+p\sin\theta)] \cdot \exp[ik\{(p\cos\theta - x')^2 + (q-y')^2\}/2d]\, dx'fy'. \quad (1)$$

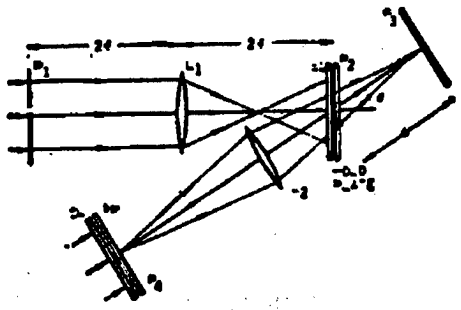


Figure 1. Fabrication of Memory Matrix for an IPA Holographic Associative Memory

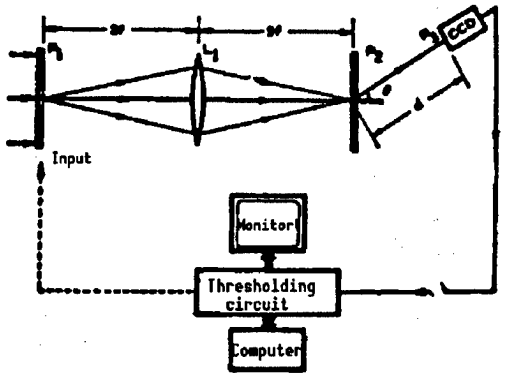


Figure 2. IPA Holographic Associative Memory

The exposure on the recording medium is proportional to

$$|\{\exp[ik(p^2+q^2)/2f] + B(p\cdot q)\}a(p,q)|^2. \quad (2)$$

Here, d is the distance between the holographic mask and plane P_3, K is the wave number of the laser, f is the focal length of the lens, θ is the angle between the optical axes of the two lenses, and D(x', y') is the diffuse scattering screen. Masks representing modes a_i and their corresponding stimulating and suppressing parts b_i are placed in positions a and b, respectively, prior to exposure. After processing, the amplitude of the transmissivity of the hologram is proportional to the amount of exposure. Terms contributing to first-order holographic diffraction include:

$$t(p,q) = \sum_i \{\exp[-ik(p^2+q^2)/2f] B_i(p,q)\} a_i(p,q). \quad (3)$$

Figure 2 shows the optical path of a holographic IPA memory. After placing the recorded Fresnel hologram on plane P_2, when the input object mask a'(x,y) is placed on plane P_1 and is illuminated by a collimated beam, the transmission holographic optical field is

$$\exp[-ik(p^2+q^2)/2f]a'(p,q)\cdot t(p,q). \quad (4)$$

On output plane P_3, the optical field can be expressed as:

$$\sum_i \iint a_i(p,q)\, a'(p,q)\, B_i(p,q) \exp[-ik(d+p\sin\theta)] \cdot \exp[ik\{(p\cos\theta - x')^2 + (q-y')^2\}/2d] \cdot dpdq. \quad (5)$$

Let $F_i(x,y)$ be the Fourier transform of $a_i(p,q)\cdot a'(p,q)$. Then, the output optical field can be written as:

$$\sum_I \iint F_I(x-x',y-y')\,D_i(x',y')\,b_I^{1/2}(x',y') \exp[ik(x^2-x'^2+y^2-y'^2)/2d]\cdot dx'dy'. \tag{6}$$

Because of the presence of the diffuse scattering screen $D_i(x', y')$, the intensity distribution can be expressed as:

$$\sum_I \iint \{a_I^2 a'^2\}\cdot\delta(x-x',\ y-y')\,b_I(x',y')\,dx'dy' = \sum_I \{a_I a'\}\,b_I(x,y)\,. \tag{7}$$

The bracket $\{a_i a'\}$ represents the inner product of a and a'. Here, the diffused light spot on the output plane is neglected. In the associative process, inner products of the input object and different parts are obtained to determine the corresponding stimulation and suppression distribution. The weighted association is the magnitude of the inner product. The magnitude of the associative matrix elements can be implemented by adjusting the diffraction efficiency of the holographic mask. In this work, different diffraction efficiencies were obtained by adjusting the exposure time to record the holographic mask.

III. Optical Implementation

Based on the above theoretical analysis, an IPA Fresnel holographic memory has been realized optically. Experimentally, three 24 x 24 pixel English letters, X, Y, and Z, were chosen as the patterns to be stored in binary form. Based on this concept, pixels can be divided into seven categories: $(\overline{X}\cap\overline{Y})\cup Z$, $(\overline{Y}\cap\overline{Z})\cup X$, $(\overline{Z}\cap\overline{X})\cup Y$, $(X\cap Y)\cap\overline{Z}$, $(Y\cap Z)\cap\overline{X}$, $(Z\cap X)\cap\overline{Y}$, $(X\cap Y\cap Z)\cap\overline{\phi}$, where $\cap$, $\cup$, and $\overline{\ }$ represent logic operations AND, OR, and NOT, respectively. ϕ represents the set of pixels none of the three patterns uses. The matrix elements with suppressing effect are negative in value. This work also adopts a commonly used scheme to provide a bias for every matrix element so that the minimum occurs at zero transmissivity in order to compute negative values optically.

In this experiment, a 30 mW He-Ne laser was used as the light source. An HP-366-1 holographic mask was used to record multiple exposures. Since there is no overlap between the light-transmitting pixels in every category, the maximum number of exposures at any point of the holographic mask is limited to one to ensure the diffraction efficiency of the mask. A two-dimensional CCD detector array is used to receive the output picture. The signal is input into a computer and the computer selects a threshold for the output picture. The associative result is then shown on the monitor. Figure 3(a) shows the picture stored in the system. Figure 3(b) is a part of the input stored by content-address finding. Figure 3(c) is the associative result with this kind of input. An IPA memory with Fresnel hologram can be used to implement pattern association with a large number of pixels. Experimentally, we have demonstrated it with 24 x 24-pixel patterns; however, this by no means is the limit. Because the resolution of the holographic mask is very high, it does

not limit the pixel number. The maximum number of pixels is determined by the input mask.

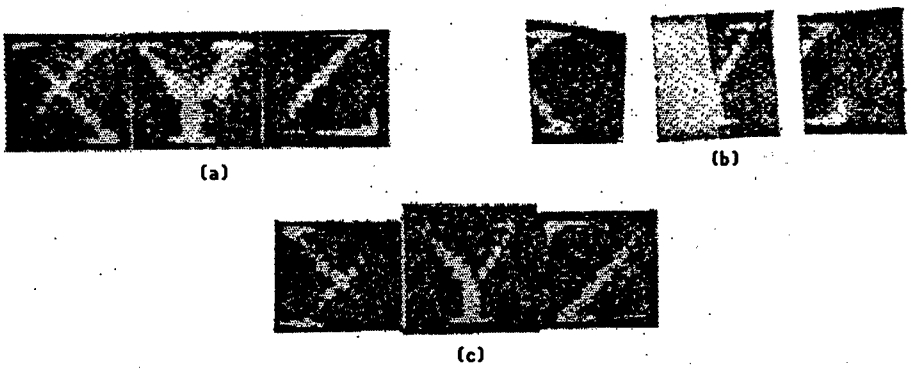


Figure 3. Stored Patterns and Recalled Outputs

Conclusions: An IPA memory with Fresnel hologram has been fabricated to increase the number of pixels to be stored in a pattern. In addition, the gray scale in the associative matrix is increased in the IPA memory to ensure high resolution of storage modes with small Hamming distance.

References

1. Taiwei Lu, Xin Xu, et al., APPL. OPT., Vol 29 No 2, Jan 1990, p 284.

Hybrid Optical-Digital System for Solving Linear Equations

92FE0151A Shanghai GUANGXUE XUEBAO [ACTA OPTICA SINICA] in Chinese Vol 11 No 10, Oct 91 pp 951-955

[Article by Li Ming [2621 2494] and Chen Yansong [7115 1484 2646] of the Institute of Physics of the Chinese Academy of Sciences, project supported by the NSFC and TWAS RG MP890-035: "Hybrid Optical-Digital System for Solving Linear Equations"; MS received 17 Dec 90, revised 19 Feb 91]

[Text] **Abstract**

A hybrid optical-digital system capable of solving a set of linear equations using an iterative method is presented. The optical part primarily consists of a single holographic mask used to carry out vector-matrix multiplication. The remaining part, including a CCD detector and a microcomputer, is for vector measurement and addition. This hybrid system was used to solve a set of four-dimensional equations. Compared to theoretical solution, the error of the experimental result is approximately 5 percent.

I. Introduction

In recent years, there are a number of studies[1-5] on optical computation of matrix multiplication. However, optical solution of linear equations is still rare. In 1982, Casasent presented a scheme for solving linear equations by using an optical vector-matrix multiplier with acousto-optic modulation and electronic feedback iteration.[2] In 1986, Caulfield, et al., presented a dual-mode optical scheme for solving linear equations using an analog and a digital optical computing system.[4] Nevertheless, they failed to present any experimental results in their papers.

It is well known that there are numerous iterative methods for solving linear equations. The most common is a simple iterative scheme. This scheme involves the inner product of a vector and a matrix and a vector addition in every iteration. The former can be done optically. As for the latter, it is more appropriate to use an electronic system. Therefore, using a hybrid optical-digital system to solve linear equations iteratively is a very logical choice.

II. Principle

Let a set of linear equations be

$$AX = B \tag{1}$$

where A and B are a two-dimensional matrix and a one-dimensional vector, respectively, and X is an unknown vector. After a primary transform of A, the iterative equation for this set of equations is

$$X^{(k+1)} = TX^{(k)} + D \tag{2}$$

where matrix T is related to A, and vector D is a function of B. The superscripts k and k+1 represent the number of iteration. Equation (2) converges only when one of the following conditions is met:

$$\left.\begin{aligned} &\max_{1\le i\le n} \sum_{j=1} |t(i,j)| < 1, \\ &\max_{1\le j\le n} \sum_{j=1} |t(i,j)| < 1, \ \sum_{i,j=1}^{n} |t_{ij}|^2 < 1, \end{aligned}\right\} \tag{3}$$

where t(i,j) is the element of matrix T in the ith row and jth column. If none of these three conditions can be satisfied, it is necessary to transform (2) further to alter matrix T and vector B until t(i,j) meets one of these conditions.

The accuracy of solving linear equations using (2) can be estimated as follows:

$$\|X^{(k+1)} - X^{(k)}\| < \epsilon, \tag{4}$$

where ϵ is an arbitrary small quantity.

The hybrid optical-digital system designed and constructed in this work is shown in Figure 1. In this system, the complex amplitude distribution of the modulating holographic mask H(m) is defined by the following equations[6-9]:

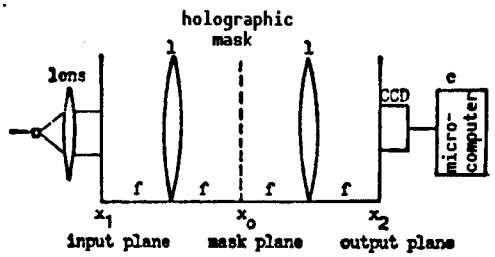


Figure 1. Scheme of Optical-Digital Computing System

$$\begin{aligned} &t(i,j) = \sum_{m} G(i,m)\,H(m)\,G(m,j), \\ &i = 1, 2, \cdots, N_2, \ m = 1, 2, \cdots, N, \\ &\qquad j = 1, 2, \cdots, N_1 \end{aligned} \tag{5}$$

In these equations, G is the Fourier transform factor; N_1, N, and N_2 are the number of sampling elements on the input plane, mask plane, and output plane, respectively. When the following equation is satisfied by optical parameters

$$(\Delta x \Delta x_1/f\lambda) = N^{-1}, \ (\Delta x \Delta x_2/f\lambda) = N_2^{-1}, \ N = N_1 N_2, \tag{6}$$

equation (5) has a unique solution. Moreover, based on unitarity, it is possible to obtain the following solution directly:

$$H(m) = \left[\sum_{i=1}^{N_2}\sum_{j=1}^{N_1} G^*(m,i)\,t(i,j)\,G^*(j,m)\right] / \sin c(r_1 m/N), \tag{7}$$

In equations (6) and (7) , Δx_1, Δx, and Δx_2 are the sampling periods of the input, holographic mask, and output phase, respectively, f is the focal length of the Fourier lens, λ is the wavelength, and r_1 is the ratio of input plane sampling element aperture to sampling period. Therefore, with a given matrix (from iteration equation (2)), the complex amplitude distribution of the holographic mask H(m) can be determined in equation (7). The design parameters of the optical system can be selected over a reasonable range based on equation (6).

III. Experiment and Results

The following is an arbitrary four-element set of linear equations:

$$\left.\begin{aligned} 38x_1 + 3x_2 + 5x_3 + 14x_4 &= 70, \quad 100x_1 + 91x_2 + 26x_3 + 48x_4 = 300,\\ 9x_1 + 6x_2 + 31x_3 + 7x_4 &= 115, \quad x_1 + 17x_2 + 30x_3 + 94x_4 = 200, \end{aligned}\right\} \tag{8}$$

This set of equations (8) is transformed into an iterative format (9):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}^{(k+1)} = -\begin{pmatrix} -0.24 & 0.06 & 0.10 & 0.28 \\ 0.24 & -0.15 & 0.16 & 0.20 \\ 0.18 & 0.12 & -0.38 & 0.14 \\ 0.01 & 0.17 & 0.30 & -0.06 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}^{(k)} + \begin{pmatrix} 1.4 \\ 1.6 \\ 2.3 \\ 2.0 \end{pmatrix}. \tag{9}$$

Comparing to equation (2), we get:

$$T = \begin{pmatrix} -0.24 & 0.06 & 0.10 & 0.28 \\ 0.24 & -0.15 & 0.16 & 0.20 \\ 0.18 & 0.12 & -0.38 & 0.14 \\ 0.01 & 0.17 & 0.30 & -0.06 \end{pmatrix}, \tag{10}$$

$$D^T = (1.4,\ 1.6,\ 2.3,\ 2.0)^T. \tag{11}$$

(T represents transformation.) After verification, matrix T meets the convergence equation (3). Equation (10) shows that $N_1 = N_2 = 4$ and $N = 16$. Substitute matrix T into (7) and let $r_1 = 0.4$, then we get the complex amplitude distribution H(m) of the holographic mask. Its numerical values are shown in Table 1.

The design parameters of the optical system are determined based on equation (6) and the existing experimental conditions. They are $\Delta x_1 = 0.2052$ mm, $\Delta x_2 = 0.8206$ mm, $\Delta x = 0.0800$ mm, $f = 415$ mm, and $\lambda = 6.328 \times 10^{-4}$mm.

Figure 1 shows the experimental scheme of the hybrid optical-digital system. The holographic mask on the confocal plane was developed using computer

generated holograms based on the data shown in Table 1. The amplitude and phase distribution extend along the x axis. Each amplitude phase value is represented by a rectangular grating extending along the y direction. If the grating period is p, starting position is y and the light transmitting aperture is Δy, then the first-order diffraction phase of the grating is $\phi = (y/p)$ and amplitude $A = [\sin(\pi\Delta y/p)/(\pi/p)]$. In this experiment, the grating period is 0.1 mm. The positioning accuracy of the grating pattern generator (GCA3600) is 0.1 μm. Therefore, the phase and amplitude of the holographic mask can be kept at one ten-thousandth.[7] The magnitude of the input vector is represented by the sampling element aperture size on the input plane. The output vector is measured by the CCD detector array. The result undergoes an A/D conversion and then is input into a microcomputer. Vector addition is done in the microcomputer.

Table 1. Amplitude-Phase Values of the Holographic Mask H(m)

m	1	2	3	4	5	6	7	8
Amplitude	0.211	1.000	0.203	0.192	0.704	0.543	0.657	0.622
Phase	1.222	-3.057	-1.923	0.822	−0.193	−0.300	2.033	−0.088
m	9	10	11	12	13	14	15	16
Amplitude	0.622	0.657	0.542	0.704	0.192	0.203	1.000	0.211
Phase	0.088	−2.033	0.300	0.193	−0.822	1.923	3.057	−1.222

The position of the holographic mask is adjusted in the following manner. On the z axis, the confocal plane is found by using a ground glass to find the maximum spot. A center line is set along the x-axis. It is lined up with the center line of sin c diffraction generated on the holographic mask plane by the input aperture. Alignment along the y-axis is not important because the grating is periodic.

Let us assume that $(1\ 1\ 1\ 1)^T$ is the initial solution $X^{(0)}$ of the set of equations. An input plate representing $X^{(0)}$ is placed on plane x_1 shown in Figure 1 and the result obtained by the optical-digital system is $X^{(1)}$. $X^{(1)}$ is then used as the input vector for the hybrid system in the second iteration. After repeated iterations, the solution of the set of linear equations can be obtained. Table 2 shows the experimental data obtained by the hybrid system and the theoretical values derived from computer-simulated iterations. They are in very good agreement.

The real solution solved by a computer is $X^r = (1.01, 0.79, 3.04, 1.01)^T$. After six iterations, the solution obtained by the hybrid optical-digital system has an error of

$$\sigma = \sqrt{\frac{1}{4}\sum_{n=1}^{4}\left[\frac{x^r(n)-x(n)}{x^r(n)}\right]^2} = 5.2\%. \qquad (12)$$

Table 2. The Iterative Solutions of the Four-Dimensional Equations Produced by the Hybrid Optical-Digital System and Theoretical Calculation

iterative time	solutions (1)		solutions (2)		solutions (3)		solutions (4)	
	E exp.	T theo.	E exp.	T theo.	E exp.	T theo.	E exp.	T theo
1	1.20	1.20	1.08	1.15	2.24	2.24	1.49	1.58
2	1.04	0.95	0.81	0.81	2.44	2.57	1.20	1.21
3	0.99	0.98	0.81	0.83	2.78	2.84	1.25	1.15
4	0.92	0.97	0.77	0.81	2.95	2.94	1.11	1.06
5	1.03	0.99	0.77	0.80	2.95	2.99	1.14	1.03
6	1.05	1.00	0.82	0.80	3.02	3.02	1.10	1.02

Due to the parallel nature of the optical system, linear equations can be quickly solved using a hybrid optical-digital system. However, since it involves an analog operation, its accuracy is somewhat limited. Nevertheless, it is useful as the front end of the scheme proposed by Caulfield for using a dual-mode optical computer to solve linear equations, because it only requires a rough solution at high speed.

The authors wish to thank Zheng Shihai [6774 1597 3189] and Li Dehua [2621 1795 5478] for their assistance.

References

1. J.W. Goodman, A.R. Dias, and L.M. Woody, OPT. LETT., Vol 2 No 1, Jan 1978, pp 1-4.

2. David Casasent, APPL. OPT., Vol 21 No 10, May 1982, pp 1859-1865.

3. David Casasent, et al., Ibid., Vol 22 No 1, Jan 1983, pp 115-124.

4. H.J. Caulfield, et al., Ibid., Vol 25, No 18, Sep 1986, pp 3128-3131.

5. Chen Yansong, Zhang Dongsheng [1728 2639 3932], and Zhu Weili [2621 0251 0448], WULI [PHYSICS], Vol 17 No 10, Oct 1988, pp 626-630; No 11, Nov 1988 pp 685-689; No 12, Dec 1988, pp 745-749.

6. Yang Guozhen [2799 0948 2823], WULI XUEBAO [ACTA PHYSICA SINICA], Vol 30 No 10, Oct 1981, pp 1340-1350.

7. Chen Yansong, Zheng Shihai, and Li Dehua, Ibid., Vol 37 No 2, Feb 1988, pp 261-268.

8. Yan-song Chen, et al., APPL. OPT., Vol 27 No 12, 15 Jun 1988, pp 2608-2611.

9. Chen Yansong, Zheng Shihai, and Li Dehua, GUANGXUE XUEBAO [ACTA OPTICA SINICA], Vol 9 No 12, Dec 1989, pp 1078-1083.

Multiple-Valued Optical Logic Implementation Via Optoelectronic Bistable Circuits

92FE0152A Shanghai ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS] in Chinese Vol 18 No 9, Sep 91 pp 691-696, 701

[Article by Liu Shutian [0491 2885 3944], Wu Jie [0702 2638], and Li Chunfei [2621 3196 7378] of the Department of Applied Physics, Harbin Institute of Technology: "Multiple-Valued Optical Logic Implementation Via Optoelectronic Bistable Circuit"; MS received 23 Jun 89]

[Text] **Abstract**

For the first time, multiple-valued optical logic has been demonstrated using a bistable laser diode (BILD) or bistable light emitting diode (BILED). Basic ternary logic functions such as COMPLEMENT, MINIMUM, MAXIMUM, CYCLE, and LITERALS are experimentally demonstrated. A triplexer circuit design is presented.

Hybrid optoelectronic logic circuits have attracted a great deal of attention lately.[1-3] A BILED-based hybrid optical multiple-valued logic circuit is presented in this paper. Experimentally, five basic ternary logic functions in POST algebra, i.e., COMPLEMENT, MINIMUM, MAXIMUM, CYCLE, and LITERALS, have been implemented. Furthermore, a ternary T-gate, or triplexer, design is shown.

Compared to binary logic, a multiple-valued logic variable carries more information. Therefore, logic operations carried out by a multiple-valued logic gate require less logic gates and connections.[4] Just as Boolean algebra is the theoretical basis for binary logic, there is a theoretical framework for multiple-valued logic. POST algebra is a common one. Considering ternary logic where R = 3, its logic variables are $x \in \{0,1,2\}$. The basic logic functions in ternary POST algebra include:

COMPLEMENT	$\overline{x} = 2 - x$,
MINIMUM	$\text{Min}(x,y) = x$, if $x \leq y$.
MAXIMUM	$\text{Max}(x,y) = x$, if $x \geq y$.

CYCLE $\quad Suc(x) = x + 1, \ (mod\ 3)$

LITERALS $\quad {}_{a}x^{nb} = \begin{cases} 2 & if\ a \le x \le b \\ 0 & elsewhere \end{cases}$

The multiple-choice switch or T-gate operation is: $\quad T(x,y,z;K) = \begin{cases} x & if\ K = 0 \\ y & if\ K = 1 \\ z & if\ K = 2 \end{cases}$

where COMPLEMENT, MINIMUM, and MAXIMUM are essentially equivalent to NOT, AND, and OR in binary logic. CYCLE can be used to implement the transformation of a logic variable. LITERALS represents the state of a logic variable. If a = b = K, then LITERALS provides the state of x at 0, 1, 2 as $J_0(x)$, $J_1(x)$, and $J_2(x)$. The T-gate can be used to select the output of a variable K at different values.

There are two factors to be considered in implementing multiple-valued logic. (1) Are there multiple input elements that can provide multiple output? (2) Can elements be conveniently connected? It is simple to connect elements in an optoelectronic circuit. We used a tristable circuit based on two BILEDs, as shown in Figure 1(a), to create a unit element for the optoelectronic ternary logic. In the figure, the LED is a red GaAsP light emitting diode, PD is a 3DU32 silicon phototriode, and T is a transistor; f_1 and f_2 are the two optical feedback coefficients for the two bistable circuits. The output amplitude is controlled by bias V and resistances R_s and R_i (i = 1,2). The thresholds to turn on $BILED_1$ and $BILED_2$ are controlled by resistors $R_{\ell 1}$ and $R_{\ell 2}$. If $R_{\ell 1} > R_{\ell 2}$, then $P_{on1} < P_{on2}$, where P_{on1} and P_{on2} are the thresholds to turn on $BILED_1$ and $BILED_2$, respectively. Due to the branching of the circuit, a tristable output is achieved with two saturated currents and the switching of $BILED_1$ and $BILED_2$ at different thresholds, as shown in Figure 1(b) [photograph not reproduced].

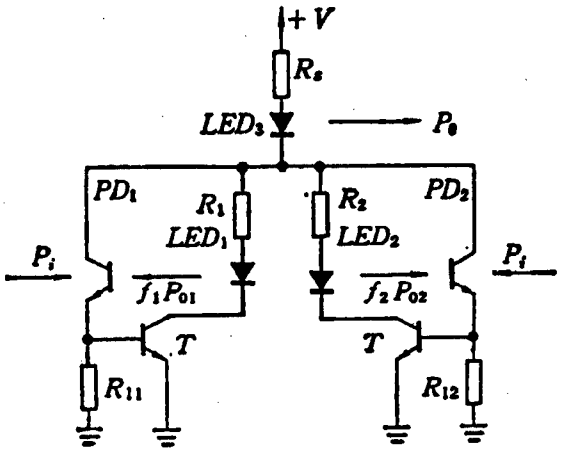


Figure 1. (a) Schematic Diagram of a Tristable Circuit or a Photo-Current Mirror; (b) Output-Input Characteristic Loop

If the three steady states are 0, 1, and 2, and the switching thresholds satisfy the following:

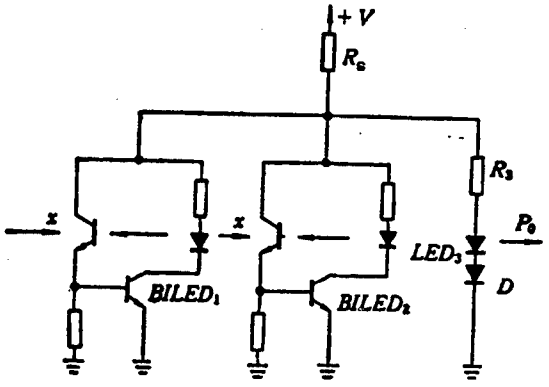
$$\begin{aligned} 0 < P_{on1} < 1 \\ 1 < P_{on2} < 2 \end{aligned} \qquad (1)$$

then the circuit can have a ternary output with respect to a ternary input in a reproducible manner. This is analogous to the current mirror in the I^2L multiple-valued logic design in electronics. The tristable circuit shown in

Figure 1(a) is called a light-current mirror, a device which plays a key role in almost all ternary logic circuits. The following is a description of optical ternary logic gates and some experimental results.

1. COMPLEMENT Gate

Figure 2(a) shows the circuit of a ternary COMPLEMENT gate, which consists of a light-current mirror and a light emitting diode LED_3. Because the output of LED_3 is opposite to that of the light-current mirror, its output P_0 appears in a downward tristable form with a triangular wave input P_i. When $x = 0$, the light-current mirror cannot be turned on and all the current flows through LED_3; consequently, $P_0 = 2$. When $x = 1$, only a part of the current flows through the light-current mirror, and output $P_0 = 1$. When $x = 2$, all the current flows through the light-current mirror and $P_0 = 0$. The input and output waveforms are shown in Figure 2(b) [photograph not reproduced].



(a)

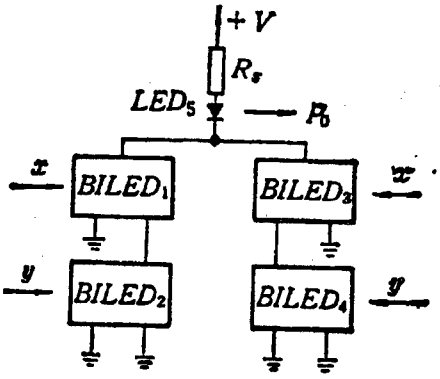
Figure 2. (a) Schematic Diagram of a Ternary COMPLEMENT Gate; (b) Input and Output Waveforms

When the amplitudes are 1 and 2, the input and output current are expressed by the LED current, which is 9.5 mA and 19.0 mA, respectively. When $P_0 = 0$, because of the added clamping diode D, the output of LED_3 is essentially held to 0. At this time, the current passing through LED_3 is approximately 0.8 mA.

2. MINIMUM Gate

A ternary MINIMUM gate consists of two light-current mirrors, as shown in Figure 3(a), where $BILED_1$ and $BILED_3$, and $BILED_2$ and $BILED_4$, form two separate light-current mirrors. Furthermore, $BILED_1$ and $BILED_2$, as well as $BILED_3$ and $BILED_4$, are also connected in series, respectively. The switching thresholds of the bistable circuits are regulated by resistors $R_{\ell i}$ ($i = 1, 2, 3, 4$):

$$0 < P_{on1} = P_{on2} < 1$$
$$1 < P_{on3} = P_{on4} < 2 \qquad (2)$$



(a)

Figure 3. (a) Block Diagram of a Ternary MINIMUM Gate; (b) Experimental Results

At this point, $BILED_1$ and $BILED_2$, and $BILED_3$ and $BILED_4$ are two separate elements to be controlled by x and y, respectively. When the input signals are 1 and 2, they are switched on simultaneously. If x and y are 1, then $BILED_1$ and $BILED_2$ are switched on at the same time and the output is $P_0 = 1$. If x and y are 2, then $BILED_3$ and $BILED_4$ are also switched on and $P_0 = 2$. Thus, it accomplishes the MINIMUM operation for x and y. Figure 3(b) [photo-graph not reproduced] shows the experimental results.

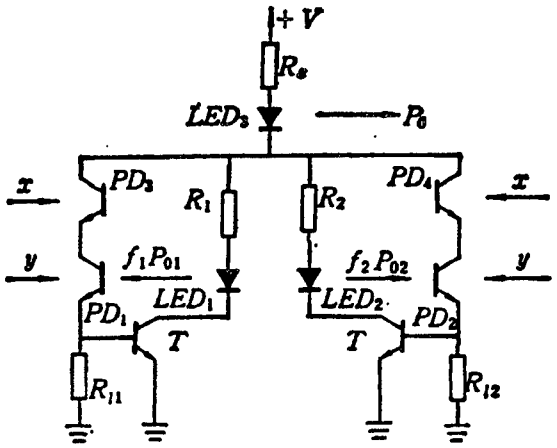


Figure 4. Circuit Diagram of the Improved Ternary MINIMUM Gate

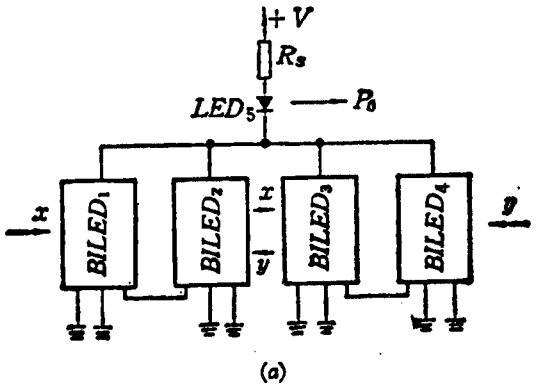


Figure 5. (a) Block Diagram of a Ternary MAXIMUM Gate;
(b) Experimental Results

In reality, a ternary MINIMUM gate can be accomplished with an improved light-current mirror to further simplify its structure, as shown in Figure 4. Two phototriodes PD_3 and PD_4 are added in series to the input end of two BILEDs. Since the photocurrent of the phototriode saturates at the minimum input light intensity, a single light-current mirror can act as a ternary MINIMUM gate as long as condition (1) is met.

3. MAXIMUM Gate

Figure 5(a) shows a schematic of a ternary MAXIMUM gate comprised of two parallel-connected light-current mirrors. The light-current mirrors are similar in construction to those described above. In addition, LED_1 and LED_2, and LED_3 and LED_4, are connected in parallel. Therefore, $BILED_1$ and $BILED_2$, and $BILED_3$ and $BILED_4$ form two separate elements to complete the MAXIMUM operation with amplitude greater than or equal to 1 and that for amplitude equal to 2, respectively. Once $BILED_1$ is switched on, $BILED_2$ is also switched on due to light feedback. $BILED_3$ and $BILED_4$ behave similarly to $BILED_1$ and $BILED_2$. If the two light-current mirrors satisfy condition (2), then only $BILED_1$ and $BILED_2$ are on when $x = 1$ or $y = 1$, and $P_0 = 1$. When x or y = 2, all the BILEDs are on, and $P_0 = 2$. Figure 5(b) [photograph not reproduced] shows the experimental results.

4. CYCLE Gate

Figure 6(a) is a schematic of a one-step-right CYCLE gate. A light-current mirror, comprised of $BILED_1$ and $BILED_2$, is connected to $BILED_3$ in parallel. A ternary signal and a constant light with amplitude 1 are simultaneously input into the system. The switching threshold of $BILED_3$ meets the following condition:

$$2 < P_{on3} < 3 \qquad (3)$$

Therefore, after the photocurrent jumps up to 2 twice in a row, $BILED_3$ is switched on to allow the current to flow from the light-current mirror through

$BILED_3$ (if the resistance of $BILED_3$ is very low relative to that of the light-current mirror). Figure 6(b) [photograph not reproduced] shows the experimental results of the one-step-right CYCLE. The input signal is x = (210) and the output is P_0 = (021).

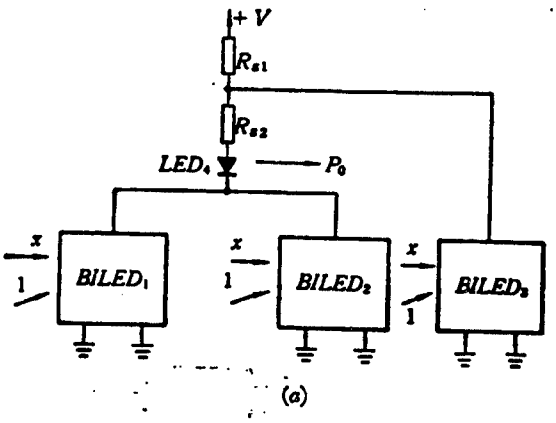


Figure 6. Schematic Diagram of a One-Step-Right CYCLE Gate

Signal x can also be cycled one step left, 0 → 2, 1 → 0, 2 → 1. Based on definition, one step left is equivalent to two steps right. Therefore, one-step-left may be accomplished by going through the gate shown in Figure 6(a) twice. Nevertheless, it may also be done with a one-step-left CYCLE gate. Figure 7(a) shows a sketch of the gate. By adding $BILED_4$ to the circuit and changing the constant light amplitude to 2, the threshold for switching $BILED_4$ on becomes:

$$3 < P_{on4} < 4 \qquad (4)$$

Hence, after $BILED_4$ is switched on, a portion of the current flows through the light-current mirror to result in P_0 = 1. This operation makes it possible to cycle x one step left.

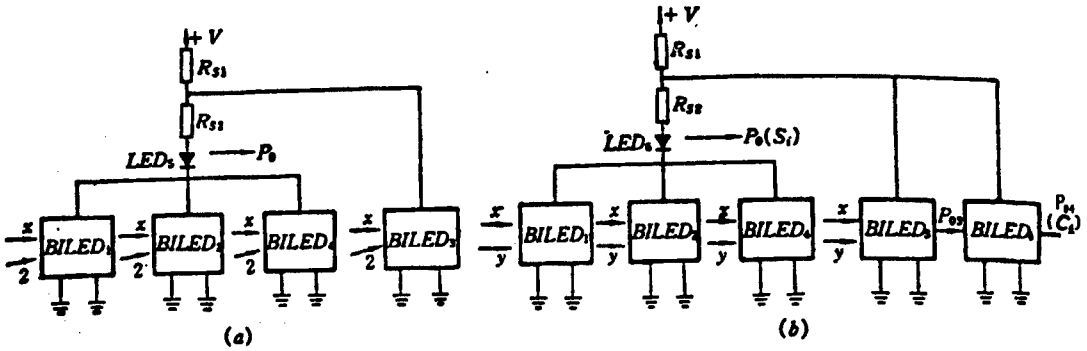


Figure 7. (a) Block Diagram of a One-Step-Left CYCLE Gate; (b) Circuit Design of a Ternary Half-Adder

It should also be pointed out that if the constant light 2 is replaced by another signal y, the circuit shown in Figure 7 is also an adder without carry. When x+y ≤ 2, the light-current mirror output is the direct addition result. When x = 1, y = 2, or x = 2, y = 1, $BILED_3$ is turned on and P_0 = 0. When x = y = 2, $BILED_4$ is switched on and P_0 = 1. Hence, the operation x AND y is realized. Please note that when $BILED_4$ is switched on, the output of $BILED_3$ is not reduced to 0. Hence, the output of $BILED_3$, P_{03}, can be used as an input for another BILED where the threshold is very low and the output amplitude is 1. This circuit can be used to output a carry. Thus, we have a ternary half adder, as shown in Figure 7(b).

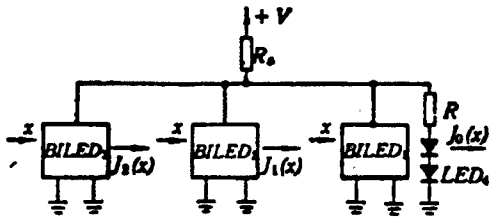


Figure 8. Block Diagram of LITERALS Gate

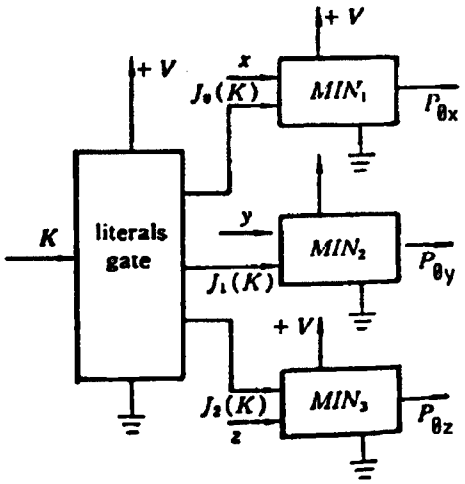


Figure 9. Block Circuit of a Multiplexer or T-Gate

5. LITERALS Gate

Functions such as $J_0(x)$, $J_1(x)$, and $J_2(x)$ can be generated by BILED circuits comprised of a binary NOT gate and an exclusive OR gate in parallel, as shown in Figure 8. $BILED_1$ and LED_4 form a binary NOT gate,[5] while $BILED_2$ and $BILED_3$ form an exclusive OR gate. The output of LED_4, $BILED_2$, and $BILED_3$ are $J_0(x)$, $J_1(x)$, and $J_2(x)$, respectively. When $x = 0$, the BILED circuit is not switched on and all the current goes through LED_4, i.e., $J_0(x) = 2$ and $J_1(x) = J_2(x) = 0$. When $x = 1$, both $BILED_1$ and $BILED_2$ are turned on and current flows through $BILED_1$ and $BILED_2$. Since there is a clamping diode in the LED_4 branch, the output of LED_4 drops to 0. In this case, $J_1(x) = 2$ and $J_0(x) = J_2(x) = 0$. When $x = 2$, the BILED is also switched on. Because the resistance of $BILED_3$ is far less than that of $BILED_2$, and there is a clamping diode in the $BILED_2$ circuit as well, $J_1(x)$ falls to 0, i.e., $J_2(x) = 2$, $J_0(x) = J_1(x) = 0$. Experimentally, the output of $BILED_3$ is far greater than 2. It is possible to normalize it to 2 by using a shunt resistor in parallel with LED_3.

6. T-Gate

Let us re-express the T-gate as follows:

$$T(x,y,z;\ K) = \text{Min}(J_0(K),\ x) + \text{Min}(J_1(K),\ y) + \text{Min}(J_2(K),\ z) \qquad (5)$$

Therefore, a ternary T-gate consists of a LITERALS gate and three MINIMUM gates. Figure 9 shows the block diagram of a T-gate. Through the LITERALS gate, a control variable K results in three functions $J_0(K)$, $J_1(K)$, and $J_2(K)$. These three functions and the three signals x, y, and z pass through three MINIMUM gates, respectively. Different variable output is obtained based on K. The simplified MINIMUM gate circuit shown in Figure 4 may be used as the ternary MINIMUM gate.

References

1. N. Bar-Chaim, K.Y. Lau, et al., PROC. SOC. PHOTO. OPT. INST. ENG., Vol 466, 1984, p 65.

2. H.D. Law., Ibid., p 69.

3. K. Okumura, Y. Ogawa, et al., IEEE J. QUANT. ELECTR, Vol QE-21, 1985, p 377.

4. S. Fukushima, T. Kurokawa, OPT. LETT., Vol 12, 1987, p 965.

5. K. Okumura, Y. Ogawa, et al., Ibid., Vol 9, 1984, p 515.

Programmable Real-Time Optoelectronic Hybrid Threshold Logic Processor

92FE0165A Shanghai ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS] in Chinese Vol 18 No 10, Oct 91 pp 760-764

[Article by Liu Shutian [0491 2885 3944], Wu Jie [0702 2638], and Li Chunfei [2621 3196 7378] of the Department of Applied Physics, Harbin Institute of Technology, Harbin 150006: "Programmable Real-Time Optoelectronic Hybrid Threshold Logic Processor"; MS received 30 Oct 89]

[Text] **Abstract**

A programmable real-time optoelectronic hybrid threshold logic processor is presented. Its applications in optical computing and optical information processing are discussed.

I. Introduction

In optical computing, the most intriguing subject is the optical neural network, which mimics the neural logic processes of the human brain.[1] Because the human brain can process logic in a highly parallel, fuzzy, multiple-valued, weighted, threshold-taking and feedback-controlled manner, it is necessary to investigate an optical logic stronger than binary vertex primitive logic, such as fuzzy logic, multiple-valued logic, and binary and multiple-valued threshold logic. In recent years, optical multiple-valued logic operations have been implemented in several optical systems.[2-4] Although Arrathoon[5] suggested using waveguides to implement optical threshold logic, very few simple optical threshold logic schemes are available. For the first time, an LED (light emitting diode) and PD (phototriode)-based optoelectronic hybrid threshold logic processor is presented in this paper. The processor can carry out real-time programmable control by optical signal to obtain two weighted and threshold values, i.e., both positive and negative weighted and threshold values. Therefore, the logic processor itself is a complete logic set. This logic processor is simple in structure, reliable, and more importantly compatible with VLSI (very large-scale integration) technology. Optoelectronic-microelectronic integration can be directly implemented on a GaAs or GaAs/Si wafer. Hence, it has a bright future in optical computing and digital optical information processing.

II. Optoelectronic Hybrid Threshold Logic Processor

Figure 1 shows the schematic diagram of a threshold logic gate. It has n binary input variables $x_i(i=1,2,\cdots\cdots,n)$. Each variable is accompanied by a weight factor $w_i(i=1,2,\cdots\cdots,n)$. There is a relation between output y and expected threshold T[6]:

$$y = \begin{cases} 1, & \text{for } \sum_{i=1}^{n} w_i x_i \geq T, \\ 0, & \text{for } \sum_{i=1}^{n} w_i x_i < T. \end{cases} \qquad (1)$$

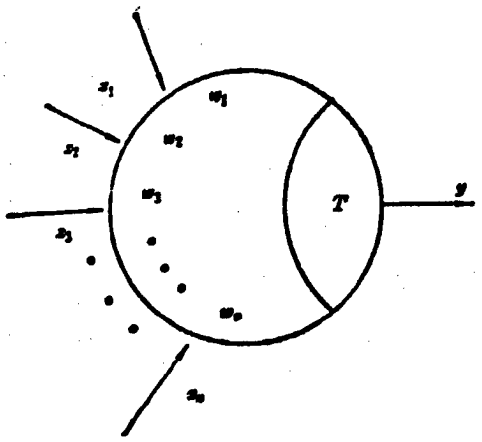


Figure 1. Schematic Diagram of Threshold Gate

Only when the weighted input variable is greater than or equal to the threshold is the output y equal to 1. Otherwise, it is 0. If a switching function $f(x)$ is a threshold logic function, i.e., it can be produced by a threshold logic gate, then there is a weight-threshold set $(w_1,w_2,\cdots,w_n;T)$ that satisfies equation (1). Mathematically, if (w_i,T) covers all real numbers, then a single threshold logic gate forms a complete logic set. This means that all logic functions can be implemented by a finite number of threshold logic gates. Furthermore, a threshold logic gate has more input capacity than a binary basic logic gate. Hence the number of logic elements and transmission lines can be significantly reduced.

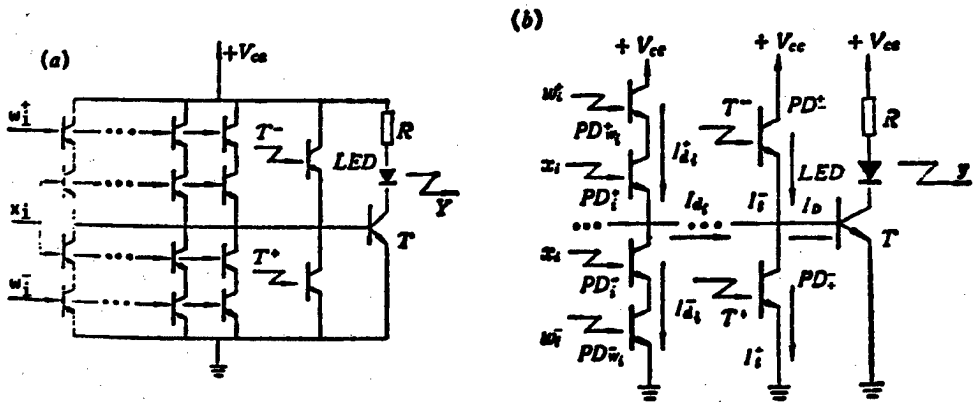


Figure 2. (a) Circuit of Optoelectronics Threshold Logic Processor; (b) Part of the Circuit of Figure 2 (a) Which Only Consists of Input x_i

Figure 2(a) shows the circuit of an optoelectronic hybrid threshold logic processor. In reality, it is a variation of a bistable LED circuit.[4] Experimentally, optical feedback was removed to increase the response speed of the logic gate. The LED output is the logic gate output y. Each weighted input end contains four silicon PDs in series. Two serve as the input of the input variable and the other two are the input of the weight signal. Each weight w_i is divided into two parts: w_i^+ and w_i^-. When x_i is positively weighted,

$w_i = w_i^+$ and $w_i^- = 0$. If x_i is negatively weighted, then $w_i = -w_i^-$ and $w_i^+ = 0$. Threshold T is also controlled by two PDs connected in series. Similarly, it has a positive and negative value. Different weight and threshold values can be obtained by varying the intensities of the weight and threshold signal. Consequently, this threshold logic processor becomes real-time programmable.

In order to illustrate how real-time weight and threshold can be implemented by the processor, let us use an example where there is only one input variable x_i, as shown in Figure 2(b). The input x_i illuminates PD_i^+ and PD_i^- and the weight control signals w_i^+ and w_i^- shine on PD_{wi}^+ and PD_{wi}^-, respectively. The current intensity I_{di} can be expressed as:

$$I_{di} = I_{di}^+ - I_{di}^- \tag{2}$$

Assuming all PDs are identical and operate in the linear amplification area, then the output photocurrent I_d is proportional to the input light intensity P_i, i.e., $I_{di} = KP_i$, where K is an optoelectronic conversion constant. Because PD_i^+ and PD_{wi}^+ (PD_i^- and PD_{wi}^-) are connected in series, I_{di}^+ (I_{di}^-) is proportional to either the x_i or w_i^+ (w_i^-), whichever has the lower optical intensity; i.e.:

$$I_{di}^{\pm} = K \min(P(x_i),\ P(w_i^{\pm})) \tag{3}$$

where min(x) is the minimum threshold function. $P(x_i)$ and $P(w_i^{\pm})$ are the intensities of x_i and $w_i^{\pm}$, respectively. The way weight and threshold values are changed is different from multiplying the amplitude of the logic variable by a factor w_i. From equation (3) we can see that $P(w_i^{\pm})$ is far less than the intensity at $x_i = 1$. Hence, $I_{di}^{\pm}$ is a function of $P(w_i^{\pm})$ only when $x_i = 1$. We have:

$$I_{di}^{\pm} = \begin{cases} KP(w_i^{\pm}), & \text{for } x_i = 1 \\ 0, & \text{for } x_i = 0 \end{cases} \tag{4}$$

Since $P(x_i)|_{x_i=1} > P(w_i^{\pm})|_{w_i^{\pm}=1}$, let $P(x_i)|_{x_i=1} = qP(w_i^{\pm})|_{w_i^{\pm}=1}$; this means that the intensity amplitude at $x_i = 1$ is evenly divided into q parts. If the weight of x_i is $\pm r (1 \le r \le q)$, then $P(w_i^{\pm}) = rP(w_i^{\pm})|_{w_i^{\pm}=1}$. For example, at $x_i = 1$, $P(x_i) = 50\ \mu W$ and at $w_i^{\pm} = 1$, $P(w_i^{\pm}) = 5\ \mu W$; then $q = 10$. If the weight of x_i is 5, then the intensity for w_i^+ is 25 μW. Of course, the value of q may be chosen arbitrarily. However, $P(w_i^{\pm})|_{w_i^{\pm}=1}$ must be greater than the saturated threshold of the PD. Usually, this threshold is very low, approximately 0.5 μW. Hence, the weight may vary over a wide range.

Let us assume that there are m positive-weight input signals and $n-m$ negative-weight input signals in the circuit of Figure 2(a). Then, the current after taking the threshold is

$$I_d = \sum_{i}^{m} K \min\{P(x_i),\ P(w_i^+)\} - \sum_{j=1}^{n-m} K \min\{P(x_j),\ P(w_j^-)\} \pm KP(T) \tag{5}$$

Here, $T = T^+$ when $\pm KP(T)$ is negative, and $T = -T^-$ when it is positive. Considering the effect of minimum saturated threshold intensity described above, the threshold intensity P(T) is adjusted so that $P(T^+) = P(T) = 0.5\ \mu W$ and $P(T^-) = P(T) + 0.5\ \mu W$. This is equivalent to adding a constant maintenance signal H where $P(H) \approx 0.5\ \mu W$ on top of PD_t^- to overcome the saturated threshold light intensity. In this case, the threshold logic output is $y = 1$ when $I_d \geq 0$. Different logic functions can be obtained by choosing different weight and threshold values.

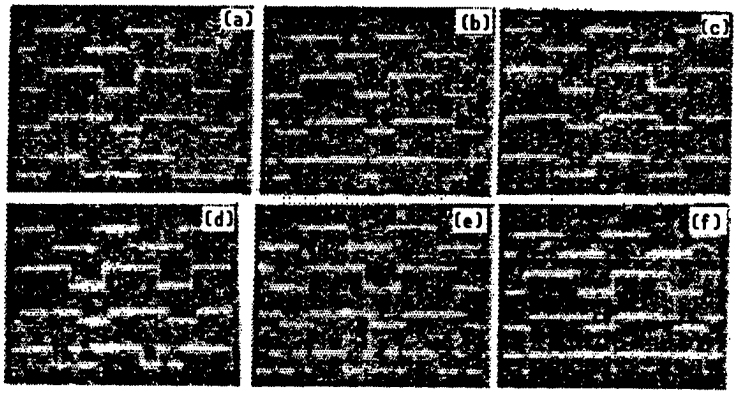


Figure 3. Experimental Results for Input Variables Number $n = 3$. The upper three waveforms indicate x_1, x_2, and x_3, respectively; the bottom waveform shows the output Y.
(a) AND; (b) OR; (c) MAJORITY; (d) $f(x) = x_1 + \bar{x}_2 x_3$;
(e) $f(x) = x_1 x_3 + \bar{x}_2(x_1+x_3)$; (f) $f(x) = \bar{x}_1 + \bar{x}_2 + x_3$

Figures 3 (a) - (f) are the experimental results with $n = 3$. The three signals were supplied by three LEDs driven by three synchronized square-wave voltage signals produced directly by the signal generator. The peak voltage is 10 V and the current-limiting resistor has a value of 220 Ω. The weight signal w_i^* and threshold T are also supplied by LEDs. By using switches and variable resistors, it is possible to change their signs and amplitudes. The voltage for the driver and the main circuit is $V_{cc} = 4.5$ V. The series resistance values corresponding to $w_i = 1$ and 2 are 1.2 kΩ and 600 Ω, respectively. Figure 3(a) and (b) shows the AND and OR logic at $T = 1$ and 3 when $w_i = 1$. Figure 3(c) shows the MAJORITY logic when $w_i = 1$ and $T = 2$. When more than two of x_1, x_2, and x_3 are 1, the output is 1. Similarly, when $w_i = -1$ and $T = -2$, we have the MINORITY logic. Obviously, the only way to get a more abundant logic function is to have different input weight values. For instance, the weight-threshold set {2,-1,1;1} can result in $f(x) = x_1 + \bar{x}_2 x_3$. With a set such as {1,-1,1;1}, we have $f(x) = x_1 x_3 + \bar{x}_2(x_1+x_3)$. With {-1,-1,1;-1} we have $f(x) = \bar{x}_1 + \bar{x}_2 + x_3$. These results are shown in Figures 3(d)-(f), and represent only a small portion of the logic functions attainable from a single threshold logic gate. It should be pointed out that only six out of eight of the possible combinations of the three input signals are shown in Figures 3(a)-(f). Two combinations, (0,1,0) and (1,0,1), are not included. However, this input signal arrangement can provide the greatest number of combinations possible. Therefore, Figures 3(a)-(f) can be considered to reflect the true values of logic operation.

A threshold logic gate can have multiple binary output values such as $y_1, y_2, \cdots, y_m$. Under identical weighting, each output end provides a different threshold. This requires the reproduction of the weight into m parts which should not interfere with each other when the threshold is applied. Figure 4 shows a logic processor with double output. The processor has two amplifier PDs with opposite output characteristics: one is an NPN triode T_1 and the other is a PNP triode T_2. LED_1 and LED_2 are the output ends. $-V_{cc}$ is the negative bias. The output of LED_2 may be 1 only with a negative weight input and greater than negative threshold value. In principle, when $\sum_{i=1}^{n} w_i x_i = T$, both outputs y_1 and y_2 should be 1. In this case, if the weight-threshold set corresponding to y_1 is $\{w_1, w_2, \cdots, w_n; T\}$, then the weight-threshold set corresponding to y_2 is $\{-w_1, -w_2, \cdots, -w_n; -T\}$. However, because of the cutoff property of T_1 and T_2, also because it cannot make correction for a threshold simultaneously, it is only possible to add H on the input ends T^- and T^+ for correction. Therefore, in a simple manner y_1 and y_2 are opposite in nature, i.e., $y_1 = \overline{y}_2$. If the threshold bias maintenance light H is removed, it requires another state $y_1 = y_2 = 0$. In the following discussion, the applications of such logic devices will be analyzed.

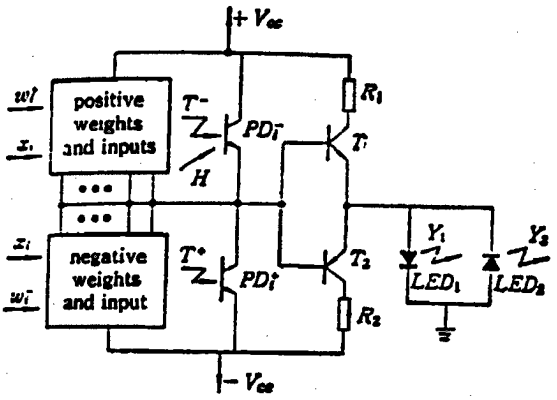


Figure 4. Schematic Diagram of Threshold Logic Processor With Double Output Y_1 and Y_2, Where $Y_1 = f(x)$ and $Y_2 = \overline{f(x)}$

III. Discussion

Threshold logic is not only an effective means to implement the logic functions in multi-variable Boolean algebra, but also can be widely used in digital graphic processing, mode recognition, classification of fuzzy functions, and neural networks. Existing optical neural network models, such as the Hopfield model,[7] are essentially based on threshold logic. Simple processes such as memory connection and learning are completed by a weighted feedback input and threshold output between neurons.[8] Therefore, threshold logic can be directly applied in neural network research. As a matter of fact, Figure 2(b) can be considered as a simple optoelectronic hybrid neural element. It can directly add and subtract weighted EXCITATORY input and weighted INHIBITORY input. Hence, this neural element can perform as a threshold gate. In comparison, the threshold gate in existing neural networks requires very complicated electronics and a computer. In addition, this neural element can have bipolar binary outputs of +1 and -1 (see Figure 4). VLSI technology and free-space optical connection technology will make this type of neural network even more attractive.

Two-dimensional parallel logic processing is also a major subject in optical computing.[9] Optical threshold logic also has some potential in this area. After connecting an OR gate in series with the threshold logic processor for the double variable input shown in Figure 4, 16 Boolean logic functions can be implemented. In general, y_1 can accomplish 14 logic functions other than XOR [i.e., exclusive OR] and XNOR [i.e., exclusive NOR]. If threshold bias is removed from the circuit of Figure 4, i.e., $y_1 = y_2 = 0$ is allowed to exist, then outputs y_1 and y_2 represent $x_1 \cdot \overline{x_2}$ and $\overline{x_1} \cdot x_2$, respectively, when $w_1 = 1$, $w_2 = -1$, and $T = 0$. When an OR gate level is added in, the OR gate's output $y = y_1 + y_2$ directly yields the XOR logic. Similarly, when $w_1 = w_2 = 1$ and $T = 1$, y directly provides XNOR logic. Such a logic processor can serve as a pixel in a two-dimensional parallel logic process. It can be directly controlled by an optical signal to implement real-time programmable space-invariant and space-variant logic.

References

1. Special Issue on Optical Computing, PROC. SOC. PHOTO-OPT. INSTRUM. ENG., Vol 1, 1988, p 693.

2. R. Arrathoon and S. Kozaitis, OPT. ENG., Vol 25, 1986, p 29.

3. M.M. Mirsalehi and T.K. Gaylord, in digest of Topical Meeting on Optical Computing, Optical Society of America, Washington, D.C., 1985, WB1.

4. Liu Shutian, Wu Jie, and Li Chunfei, ZHONGGUO JIGUANG [CHINESE JOURNAL OF LASERS] in Chinese Vol 18, 1991 p 691 [see "Multiple-Valued Optical Logic Implementation Via Optoelectronic Bistable Circuits," translated in full above].

5. R. Arrathoon and M.H. Hosson, OPT. LETT., Vol 9, 1984, p 143.

6. S.C. Lee, "Modern Switching Theory and Digital Design," Prentice-Hall, Inc., Chapter 4, 1978, p 115.

7. J.J. Hopfield, PROC. NATL. ACAD. SCI., USA79, 1982, p 2554.

8. D. Psatis and N. Farhat, OPT. LETT., Vol 10, 1985, p 98.

9. S. Fukushima and T. Kurokawa, Ibid., Vol 12, 1987, p 965.

- END -